



CAC 68

CARTOGRAPHIE DES ACTIVITES DE RECHERCHE DU LABORATOIRE COGIT 2006-2008

PREMIERE PARTIE

VUE GENERALE SUR LES RECHERCHES DU COGIT

Présentation	4
Les recherches 2006-2008	4
Composition de l'équipe	5
Publications et Site web.....	7
Projets, Collaborations, Animation scientifique et Formation	8
GeOxygène	10

DEUXIEME PARTIE

LES RECHERCHES DU COGIT PAR THEME

Thème 1 Intégration des données	
Présentation générale	12
Présentation par chercheurs	15
Thème 2 Généralisation	
Présentation générale	20
Présentation par chercheurs	26
Thème 3 Sémiologie	
Présentation générale	30
Présentation par chercheurs	34
Thème 4 Accès aux services et aux données géographiques	
Présentation générale	37
Présentation par chercheurs	42
Thème 5 Analyse de l'espace géographique et de ses dynamiques	
Présentation générale	49
Présentation par chercheurs.....	55

TROISIEME PARTIE

PUBLICATIONS ET COMMUNICATIONS

Thèses.....	62
Livre, chapitres de Livre, Actes de conférences	62
Revue	62
Conférences A	63
Conférences B	64
Autres communications	66
Rapports techniques et rapports de stage	67

PREMIERE PARTIE

**VUE GENERALE
SUR LES RECHERCHES DU COGIT
2006-2008**

PRESENTATION

Partie rédigée par Anne Ruas

Contexte

Le laboratoire COGIT est l'un des quatre laboratoires de recherche de l'Institut Géographique National. Il a été créé en 1988, en même temps que le service de la recherche de l'IGN. Les axes de recherche du laboratoire sont décrits dans le schéma directeur de la recherche et du développement de l'IGN (SDRD), rédigé par les laboratoires de recherche et la direction technique de l'IGN. Le SDRD est validé par la direction de l'IGN après avis auprès du Comité Scientifique et Technique (CST) de l'IGN nommé par le gouvernement (Arrêté du 8 Mars 2006).

Ce rapport d'activité propose une Cartographie des Activités de recherche du Cogit sur la période 2006–2008 (CAC68). Il s'inscrit dans le SDRD 2006-2010 écrit en 2005. Il suit le bilan précédent (CAC105) qui décrivait les recherches du COGIT des années 2001-2005.

Les bases de données topographiques vectorielles, produits et informations dérivés

Par rapport aux recherches générales en Géomatique, le COGIT reste proche du domaine d'expertise de l'IGN à savoir la constitution, la gestion et l'exploitation de bases de données topographiques. Celles-ci s'attachent à décrire précisément la position géographique d'entités du monde réel à l'aide d'objets dont la géométrie est décrite par des points, lignes et surfaces. La partie thématique est représentée sous forme d'attributs. Ces BD décrivent les constructions humaines (bâtiments, routes, voies ferrées, etc.) ainsi que le paysage (les rivières, les lacs, les forêts, le relief, etc.). Une base de données est généralement conçue pour répondre à un ensemble de besoins pressentis lors de sa constitution. La réponse à un type de besoin précis entraîne le plus souvent des transformations de l'information initiale. Une fois la base disponible, elle devient une source d'informations qui est destinée en premier lieu à l'usage pressenti mais comme toute source d'informations elle peut être modifiée (abstraite, réorganisée, complétée) pour s'adapter à de nouveaux besoins.

Ainsi, les recherches du COGIT sur les bases de données topographiques traitent non seulement de la représentation des données (leur modélisation, leur mise en cohérence, leur enrichissement), mais aussi de l'automatisation de la fabrication de produits dérivés de ces bases qu'il s'agisse de cartes ou d'extraits de bases. Dans ce contexte, l'objectif principal du laboratoire est d'identifier et de formaliser les connaissances, les modèles et les interfaces qui permettent de répondre au mieux aux besoins en informations géographiques qui sont dérivées de bases de données topographiques vectorielles. La formalisation concerne donc les données, les traitements appliqués sur ces données, et l'accès aux données et aux traitements.

LES THEMES DE RECHERCHES

Le COGIT a réalisé des recherches selon les cinq axes suivants :

Axe 1 : Intégration des bases de données, gestion multi-résolution

Objectifs : Intégrer, gérer et visualiser des données ayant des niveaux de détails comparables ou différents (en utilisant alors la représentation multiple).

Axes de recherche : Mise en relation de schémas de données, en regardant spécifiquement le problème de l'intégration de propagation de mise à jour dans des bases de données hétérogènes (typiquement données métier et la BDUNI de l'IGN). Construction d'un outil de représentation des spécifications de saisie. Construction d'un outil de représentation, de manipulation et de visualisation des BD à représentation multiple. Mise en cohérence, gestion des raccords et appariement; construction d'une base par complémentarité maximale.

Axe 2 : Généralisation

Objectifs : Automatiser la fabrication de cartes ou de bases de données avec changement de niveaux de détails.

Axes de recherche : Poursuite de l'automatisation des procédures de généralisation de données géographiques. Développement de la généralisation non cartographique (de BD à BD). Généralisation des moyennes échelles avec prise en compte spécifique des réseaux.

Axe 3 : Cartographie

Objectifs : Développer des méthodes permettant d'améliorer la qualité des légendes des produits cartographiques et leur adaptation (« carte à la carte ») en fonction de contraintes spécifiques à l'utilisateur.

Axes de recherche : Amélioration des contrastes de couleurs. Réalisation de légendes à la carte. Analyse des contrastes visuels.

Axe 4 : Accès aux services et aux données géographiques

Objectifs : Faciliter l'accès aux services de l'information géographique en développant des procédés permettant à l'utilisateur, non nécessairement spécialiste, d'accéder de manière intuitive à des données et des traitements géographiques éventuellement complexes.

Axes de recherche : Conception et développement d'une plateforme de module de consultation. Structuration de données. Conception d'interfaces « homme/machine » multimodales notamment pour la conception de légendes. Description des traitements géographiques et de leur mode d'emploi ; déclenchement distant de requêtes d'extraction, de simplification et de structuration sur les BD

Axe 5 : Analyse de l'espace géographique et de ses dynamiques

Objectifs : Participer activement au meilleur usage des données topographiques dans des applications utiles à la société.

Axes de recherche : Phénomènes naturels et les risques qu'ils engendrent ; caractérisation du relief et mise en cohérence de données 2D, 2D½ et 3D ; développement de partenariats sur l'aspect aléa du risque. Réseaux de transport et leur description, modélisation multi niveaux de ces réseaux. Espaces urbains et leur dynamique, analyse des phénomènes émergents.

COMPOSITION DE L'EQUIPE

Le COGIT est composé de chercheurs (essentiellement des ingénieurs de l'IGN qui ont soutenu une thèse de doctorat), d'ingénieurs d'étude de l'IGN et de doctorants qui sont soit des ingénieurs de l'IGN en thèse au COGIT soit des doctorants sur bourse de thèse IGN. En 2007 un chargé de recherche du ministère a été embauché pour diversifier le recrutement. Les recherches sont structurées selon les axes présentés précédemment. Elles sont encadrées par des chercheurs du COGIT.

A l'exception d'une thèse de doctorat (ED La Rochelle), les thèses sont inscrites soit à l'école doctorale ICMS (Information, Communication, Modélisation, Simulation) de Paris-Est (5 fin 2008) ou à l'école doctorale de Géographie de Paris (2 fin 2008). Les thèses inscrites à Paris Est sont dirigées par Anne Ruas.

Responsable du Laboratoire : A. Ruas (IPC¹, HDR en 2004)

Chercheurs IGN :

En activité au COGIT en Janvier 2009 : B. Bucher (IPC, Dr en 2002); C. Domingues (ITPE², Dr en 2001); C. Duchêne (ITGCE³, Dr en 2004) ; J. Gaffuri (ITGCE, Dr en 2008), JF Gleyze (IPC, Dr en 2005) ; S. Mustière (ITGCE, Dr en 2001), J. Perret (CR, Dr en 2007)

Départ du COGIT entre 2006 et fin 2008: S. Balley (ITGCE, Dr en 2007), O. Bonin (ITGCE, Dr en 2002) P Bordin (IPC, Dr en 2006),

¹ IPC : Ingénieur des Ponts et Chaussées

² ITPE : Ingénieur des Travaux Publics de L'Etat

³ ITGCE : Ingénieur des Travaux Géographiques et Cartographiques de l'Etat

Ingénieurs d'étude IGN

En activité au COGIT en Janvier 2009 : M. Brasebin (ITGCE); L. Jolivet (ITGCE), J Renard (ITGCE)

Doctorants IGN :

Inscrit en thèse en Janvier 2009 : E. Grosso (ITGCE), G. Touya (ITGCE), N. Abadie (ITGCE); E. Buard (ITGCE), F. Dhee.

Doctorants sous contrat IGN :

Thèses en cours en janvier 2009: S. Christophe ; E Mermet ; JF Girres

Thèses soutenues entre 2006 et 2008 : Y. AbdElKader(2006) ; E. Chesneau (2006) ; B. Poupeau (2008) ; AM. Olteanu (2008) ; P. Taillandier (2008)

Ingénieur d'étude sous contrat IGN 2006-2008 : Adrien Paget

Doctorants sous contrat non IGN

Thèse soutenue: C. Pierkot (CIFRE EADS, 2008)

Chercheurs Invités

Stefan Steiniger (Doctorant, Université de Zürich), 1 mois, octobre 2006

Teodor Foerster (Doctorant, ITC Enschede), 1 mois, janvier 2007

Byron Nakos (Professeur, Université d'Athènes), 6 mois en 2007

Henrike Firkowski (Professeur, Université de Curitiba, Brésil), 1 an en 2007, post-doc

Magali Valdeperez-Pastor (Stagiaire, IGN Espagne), 2 mois fin 2007

Ozgür Dogru (Doctorant, Université d'Istanbul), 1 an en 2008

On notera que tous les chercheurs ou doctorants invités sont venus travailler sur le thème de la généralisation

Stagiaires : sur la période 2006-2008, le COGIT a accueilli 16 stagiaires

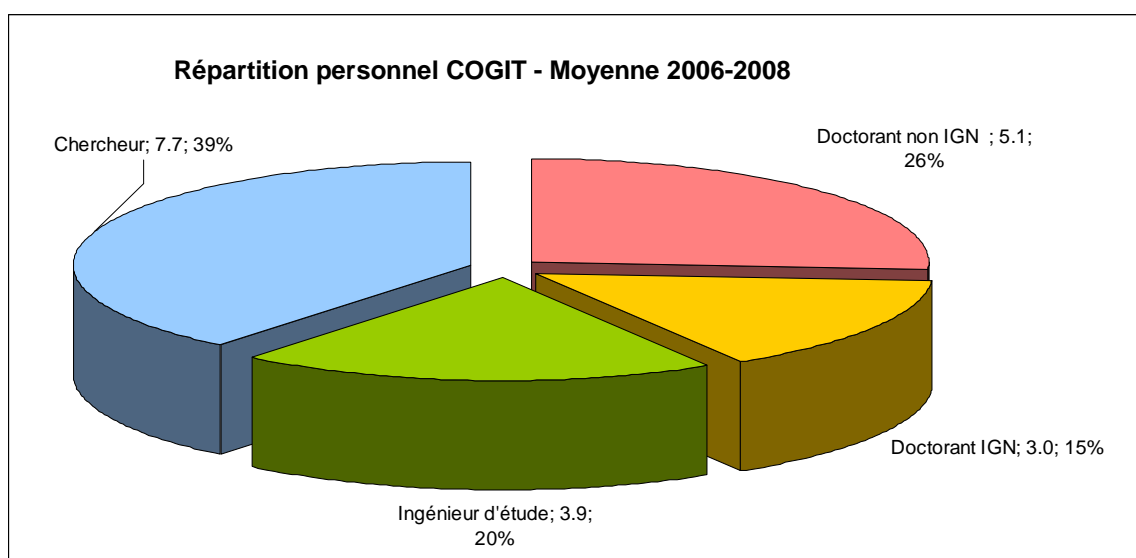
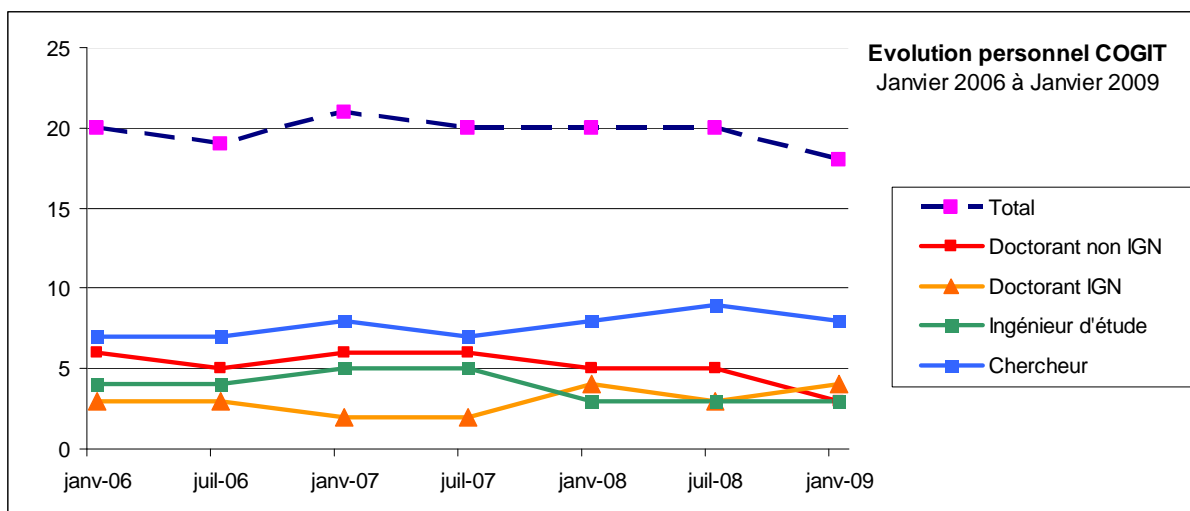
Personnel par action de recherche en Janvier 2009 :

Personnel par action de recherche - Janvier 2009		Services	Intégration	Généralisation	Espace et dynamique	Sémiologie
Encadrement		Bucher	Mustière	Duchêne	Gleyze	Ruas
Chercheurs		Domingues		Gaffuri	Perret	
Doctorants	3 ^{ème} 2 ^{ème} 1 ^{ère}	Christophe	Abadie; Grosso	Touya	Mermet Buard; Girres	Dhée*
Ingénieurs d'étude		Jolivet		Renard	Brasebin	
Total		4	3	4	6	1,3

* ENSG

Quantité et Répartition par profil

Sur les trois années en moyenne le COGIT a été composé de 19,7 agents, dont 39% de chercheurs, 41% de doctorants et 20% d'ingénieurs d'étude. L'âge moyen est de 30 ans et la parité femme-homme est assurée (voir tableaux et graphiques ci-dessous).



Age et Parité

	janv-06	juil-06	janv-07	juil-07	janv-08	juil-08	janv-09	Moyenne 2006-2008	En 2001
Nbre personne	20	19	21	20	20	20	18	19.7	20.5
Age moyen	29.8	30.5	30.3	30.4	30.0	30.5	30.7	30.3	30.3
Parité Femme/Homme	1.2	1.1	1.2	1.1	1.0	1.0	0.8	1.1	0.34

PUBLICATIONS ET SITE WEB

Les publications sont présentées en partie 3 de ce rapport et synthétisées dans le tableau ci-dessous (les chiffres pour la période 2001-2005 sont donnés en guise de comparaison)

Les publications sont structurées en 5 catégories : 1/ les livres, chapitres de livre ou éditions d'ouvrage, 2/ les revues avec comité de sélection sur articles long, 3/ les conférences rang A avec une sélection sur article long et comité de lecture, 4/ les conférences de rang B avec une sélection sur résumé et comité de lecture et 5/ les autres communications.

Le nombre de publications annuelles en revues et en conférences de rang A et B par chercheur (docteurs et doctorants hors première année de thèse) est de 2,1 alors qu'il était de 1,3 sur la période 2001-2005.

Il y a 3 thèses soutenues en moyenne par an. Elisabeth Chesneau et Nils Gesbert ont eu la première et deuxième place des meilleures thèses en Géomatique du GdRSigma sur la période 2002-2006. Sandrine Balley a eu le deuxième prix de la meilleure thèse du GdR en 2007.

	2006	2007	2008	Total sur les 3 ans	moyenne / an	ref 2001-2005 (5 ans)
HDR	0	0	0	0		1
Thèses	3	1	5	9		14
Livre	0	0	0	0		2
Chapitre de Livre	0	3	1	4		11
Edition d'ouvrage	0	1	1	2		1
Revue	2	3	5	10	3,3	3,4
Conférence RgA	5	7	6	18	6,0	5,8
Conférence RgB	18	18	14	50	17	9,6
Total Revue + Conf	25	28	25	78	26	18,8
Nb Personnel	19,5	20,5	20		20	19,6
Publication / Personne	1,3	1,4	1,3		1,3	1,0
Nb Chercheurs*	12,0	13,0	13,0		12,7	14,6
Publication / chercheur	2,1	2,2	1,9		2,1	1,3
Autres communications	15	16	5	36	12,0	12

* = Docteurs et doctorants après la première année de thèse

Site web

Le site web du COGIT (<http://recherche.ign.fr/labos/cogit>) a été complètement reconçu en 2003 par Sandrine Balley puis mis à jour par Elodie Buard. Une version anglaise est disponible.

Les données relatives aux chercheurs, aux axes de recherche et aux publications sont stockées dans un SGBD (MySQL), ce qui facilite les mises à jour du site. La structure du site et les programmes associés ont été fournis aux autres laboratoires de l'IGN.

Les thèses et HDR sont en ligne: <http://recherche.ign.fr/labos/cogit/telechargementCOGIT.php>

PROJETS, COLLABORATIONS, ANIMATIONS SCIENTIFIQUES, FORMATION

Dans les années 2005-2007, le COGIT a participé à l'ANR **Tadorne** (ACI Masse de données) (<http://ufrsciencetech.u-bourgogne.fr/~gadavid/tadorne/>) sur le tatouage de données géographiques en collaboration avec le laboratoire CEDRIC du CNAM (David Gross-Amblard, Julien Lafaye), le GREYC de Caen (Jacques Madelaine, Jean-Marie Le Bars), le LAMSADE de Paris-Dauphine (Cristine Bazgan). Plusieurs algorithmes ont été conçus dont un par le COGIT sur le tatouage des bâtiments (voir le stage de Jean Beguec et publications résultantes).

Deux projets ANR GeOnto et GeOpenSim ont démarré en 2008. GeOpenSim est dirigé par le COGIT.

- **GeOnto** (<http://geonto.lri.fr/>) vise à faciliter l'intégration de documents géographiques par appariement des concepts (ontologies) sous-jacents. Les partenaires sont 1/ L'équipe IASI/Gemo du LRI (Laboratoire de Recherche en Informatique) sous la tutelle du CNRS de l'université Paris-Sud/Orsay et de l'INRIA Futurs. Cette équipe, dirigée par Chantal Reynaud, est spécialiste en particulier d'intégration sémantique de données. 2/ L'équipe IC3 de l'IRIT (Institut de Recherche en Informatique de Toulouse), UMR commune au CNRS, à l'Institut National Polytechnique de Toulouse, à l'Université Paul Sabatier et à l'Université des Sciences Sociales Toulouse 1. Cette équipe, dirigée par Nathalie Aussenac-Gilles, est spécialiste en particulier de la construction d'ontologies et de terminologies à partir de textes. Et 3/ L'équipe Document Electronique, Sémantique et Interaction (DESI) du laboratoire d'informatique de l'université de Pau. Cette équipe, dirigée par Mauro Gaio, est en particulier spécialiste de la modélisation et du traitement des contenus documentaires à connotation géographique

- **GeOpenSim** (<http://geopensim.ign.fr/>) vise à construire un module OpenSource de simulation des évolutions des structures urbaines en se basant sur l'analyse des bases de données topographiques de différentes temporalités. Les partenaires du projet sont 1/ le laboratoire Image et Ville de Strasbourg dirigé par Christiane Weber, spécialisé sur les analyses urbaines, 2/ le laboratoire informatique LSIIT de Strasbourg dirigé par Pierre Gancarski, spécialisé en Apprentissage et 3/ la chercheuse Annabelle Mas de l'université d'Orléans spécialiste en analyse urbaine.

Le COGIT participe aux projets **TerraNumérica** et **TerraMagna** du pôle de compétitivité CapDigital pour des tâches de conception suivantes :

- Conception de services web (TerraNumérica),
- Conception d'un algorithme de simplification de bâtiments 3D (TerraNumérica),
- Conception d'algorithmes de manipulation de structures 3D (TerraMagna),
- Conception d'un langage et d'outils de représentation de règles d'urbanismes (TerraMagna).

Les chercheurs du COGIT participent aux groupes de travail du GdRSigma (modélisation, sémiologie, mutualisation) et du GdRE S4 (Jean-François Gleyze sur l'analyse des réseaux).

Animation de la recherche

En termes d'animation de recherche, certains chercheurs du laboratoire ont des responsabilités nationales ou internationales :

Anne Ruas :

- est vice-présidente de l'ACI (<http://cartography.tuwien.ac.at/ica/>) et du CFC,
- a co-organisé avec Chris Gold la conférence Spatial Data Handling à Montpellier en 2008 (Springer), et est responsable de l'organisation de la conférence de l'ACI à Paris en 2011,
- est membre du bureau du GdRSigma et Magis et co-responsable d'un groupe de travail sur la modélisation (Sigma) et sur les dynamiques spatiales (Magis),
- a réalisé plusieurs évaluations pour l'AERES en 2008, participe au jury du concours interne d'admission dans le corps des IPC et à un jury de recrutement d'un maître de conférences en SIG, et est membre du conseil scientifique de l'EIVP et de l'ESGT.

Sébastien Mustière :

- est responsable d'une commission de l'ACI (<http://ica.ign.fr/>) et d'une commission du CFC,
- a co-organisé avec William Mackaness, Jati Stoter et Peter Van Oosterom un workshop en généralisation à Montpellier en Juin 2008,
- est expert INSPIRE sur le thème de la toponymie.

Bénédicte Bucher :

- est co-responsable d'un groupe de travail du GdRSigma et Magis sur la Mutualisation,
- a participé aux comités de programme de SAGEO et de SDH.

Cécile Duchêne :

- a participé au jury d'admission au concours interne des ITGCE,
- a participé au comité de sélection des workshops de la commission de généralisation de l'ACI et au test d'EuroSDR (www.eurocdr.net) sur la comparaison de plateformes de généralisation.

Formation

Les chercheurs du COGIT donnent environ 500 heures de cours par an dont 300 à l'ENSG, 80 à Paris-Est et le reste auprès de différents formations en géomatique dont les masters SIIG3T et ATEG.

Le COGIT est responsable de 6 modules d'enseignement: 2 modules dans le Master SIG ENSG-Paris-Est (*Données et SIG* A. Ruas ; *SIG et Internet* B. Bucher), 1 module au cycle des ingénieurs IT2 de l'ENSG (*Analyse Spatiale* A. Ruas), 1 Module au master IST de Paris-Est (*Données Géographiques* A. Ruas), 1 Module au Master GeoRisque de Paris-Est (*Données Géographiques* A. Ruas puis J Gaffuri), 1 Module Master Carthageo Paris1-Paris7-ENSG (*Modélisation de données* Sébastien Mustière).

GEOXYGENE

Les chercheurs du COGIT utilisent deux SIG : *Gothic/Clarity* de la compagnie 1Spatial pour les recherches en Généralisation et *GeOxygène* pour les autres recherches.

La plateforme GeOxygène a été conçue au COGIT depuis 2000 à partir de développements propres en Java et de bibliothèques OpenSource. Le modèle objet s'inspire au maximum des normes ISO. Les données sont stockées sur le SGBD libre PostGIS en remplacement d'Oracle initialement utilisé. Une partie du code de GeOxygène est déposée régulièrement en OpenSource sur le site sourceforge (oxygene-project.sourceforge.net/) et le reste (appelé ci-dessous code interne) est en cours de conception au COGIT. Le dépôt en OpenSource nous permet de faciliter les collaborations avec d'autres laboratoires dont ceux qui accueillent les chercheurs qui ont fait leur thèse au COGIT.

GeOxygène est le logiciel d'accueil du projet GeOpenSim. Les développements seront déposés en OpenSource à la fin du projet et viendront enrichir la plateforme.

Sur la période 2007-2008 :

- un module d'appariement a été déposé en OpenSource. Il a été repris par l'IGN Belge pour la production en 2009 de l'appariement de ses données au 1/10.000 et au 1/50.000 en vue de faciliter leur mise à jour. Ce module est aussi en cours de test à l'Ordnance Survey.
- des formations gratuites et ouvertes ont été données aux chercheurs du GdRSigma, aux chercheurs de GeOpenSim, à l'ENSG et à l'IGN-Belge, ainsi que de l'aide en ligne auprès de l'EPFL, l'IGN-Belge et l'Ordnance Survey.
- plusieurs améliorations ont été faites sur le code interne de GeOxygène dont les suivantes :
 - codage du lien entre le schéma de données et les métadonnées à partir du travail de thèse de Sandrine Balley,
 - chargement de données 3D, amélioration de la gestion du stockage en shape,
 - préparation de deux modules pour un dépôt en 2009 :
 - o conception d'un module de visualisation et de manipulation de données 3D à partir entre autre du travail de thèse de Benoit Poupeau
 - o portage et amélioration du module d'amélioration des couleurs des légendes à partir du travail de thèse d'Elisabeth Chesneau

DEUXIEME PARTIE

**LES RECHERCHES DU COGIT PAR THEME
2006-2008**

THEME 1 : INTEGRATION DE DONNEES

Rédigé par Sébastien Mustière, responsable de l'action de recherche sur l'intégration.

OBJECTIFS ET CONTEXTE

Avec d'une part l'essor des technologies de l'information, et d'autre part le développement des techniques de localisation spatiale, les données géographiques sont de plus en plus nombreuses et variées. Or, chacune de ces bases de données géographiques reflète une conceptualisation particulière du monde réel, propre à son producteur. Aussi, d'une base de données à l'autre, un même phénomène du monde réel pourra avoir différentes descriptions et interprétations, difficulté qui peut être qualifiée « d'hétérogénéité sémantique » des données. S'il ne s'agit pas d'une caractéristique exclusive des données géographiques, cette hétérogénéité est particulièrement vérifiée dans ce domaine, en raison de la complexité du monde géographique représenté.

L'hétérogénéité des informations pose problème à l'utilisateur qui souhaite analyser conjointement diverses données, qu'elles soient multi-thèmes, multi-échelles, multi-vues ou transfrontalières. Elle pose aussi problème au producteur de données qui doit gérer la cohérence entre ses données et éviter les tâches redondantes pour leurs mises à jour. Elle pose également problème aux producteurs qui souhaitent échanger de l'information pour enrichir mutuellement leurs données.

Par ailleurs, la gestion de l'hétérogénéité est en train de devenir une obligation pour les producteurs de données publiques, à travers la directive Européenne INSPIRE, qui impose en particulier « qu'il soit possible de combiner de manière cohérente des données géographiques tirées de différentes sources dans la Communauté et de les partager entre plusieurs utilisateurs et applications », et plus précisément qui demande que « les règles de mise en oeuvre [de la directive soient] conçues pour assurer la cohérence entre les éléments d'information qui concernent le même lieu ou entre les éléments d'information qui concernent le même objet représenté à différentes échelles ».

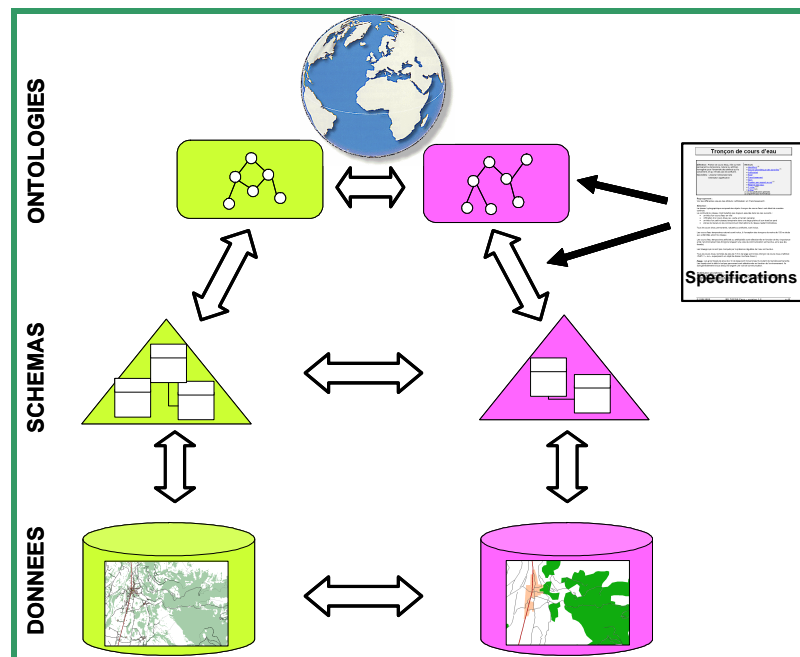
Il est donc nécessaire d'appréhender l'hétérogénéité des données ou, en d'autres termes, d'intégrer des données au moins virtuellement afin d'explicitier leurs relations. Le principe même des bases de données est de séparer les données elles-mêmes des schémas qui les décrivent. Il en découle que la problématique de l'intégration peut se décomposer en deux sous-problèmes. Le premier est de comprendre comment les données s'apparient et éventuellement de les recaler les unes sur les autres. Le deuxième est de comprendre comment les schémas de bases de données se correspondent et éventuellement de créer de nouveaux schémas pour avoir une description harmonisée de l'information.

APPROCHES

Dans ce contexte, des travaux de recherche sont à effectuer à plusieurs niveaux pour intégrer des données géographiques, de plus en plus nombreuses et hétérogènes. L'approche générale du laboratoire COGIT sur ce sujet, illustrée sur la figure suivante et détaillée ci-dessous, repose sur deux principes : premièrement, il est nécessaire de travailler à la fois au niveau des schémas et des données car ces niveaux sont inter-dépendants et, deuxièmement, les spécifications des bases de données sont des documents riches et utiles pour en comprendre la sémantique et savoir les prendre en compte facilite l'intégration de données.

Le premier niveau de nos recherches est celui des données. Si différents travaux sur l'appariement de données ont été effectués pour répondre à des problèmes précis, à l'IGN comme ailleurs, on constate un manque de généralité des outils proposés. Pour arriver à une certaine généralité, nous pensons qu'il est nécessaire de mettre au point un processus qui combine de manière efficace de nombreux critères d'appariement, même reposant sur des informations imparfaites. La gestion de cette imperfection pour l'appariement a fait l'objet de la thèse d'Ana-Maria Olteanu-Raimond de 2005 à 2008 au COGIT. Par ailleurs, pour des raisons techniques, philosophiques et politiques, nous pensons que les outils d'appariement doivent être fournis à travers un réseau sous la forme de services web qui ne nécessitent pas de regrouper les données à apparier au sein d'un même support physique. Apparier à la demande

et en fonction de chaque besoin est une approche qui nous semble plus réaliste et souple que la quête d'un appariement unique, fait 'une fois pour toute'. Différentes solutions sont envisageables, parmi elles le recours à des services web constitue une solution a priori pertinente. Des études autour des services web ont ainsi été réalisées au COGIT.



Approche générale de l'intégration au COGIT

Le deuxième niveau de nos recherches est celui des schémas. Pour résoudre les divers problèmes liés à l'hétérogénéité sémantique, un consensus se fait globalement autour de la nécessité de disposer d'ontologies, c'est-à-dire d'outils permettant de spécifier sans ambiguïté la sémantique de systèmes terminologiques donnés. Celles-ci sont encore peu nombreuses dans le domaine de la géographie, même si l'on recense quelques travaux de mise en place d'ontologies de haut niveau ou d'ontologies de domaine comme *Towntology* et *Fodomust* en France. De façon générale, la création d'une seule et même ontologie du domaine géographique permettant l'intégration de tous les points de vue existants reste difficilement réalisable, voire peu souhaitable. Il nous semble donc important que chaque communauté d'acteurs de l'information géographique constitue sa propre ontologie du domaine, reflétant son point de vue sur le monde, et explicitant la sémantique des concepts partagés par ses membres. Dans ce cadre, nous effectuons des recherches sur la création et l'alignement d'ontologies géographiques représentant le point de vue des topographes. Nous proposons d'initier cette tâche en exploitant le contenu des spécifications des bases de données de l'IGN.

Par ailleurs, ces mêmes spécifications explicitent de manière informelle mais très précise la sémantique des éléments des schémas de bases de données. Nous pensons que formaliser cette information riche est utile voire nécessaire pour atteindre un fort degré d'intégration des données et une certaine automatisation des processus. Pour cela, un formalisme a été défini dans nos recherches passées (cf. thèse de Nils Gesbert soutenue en 2005). De 2006 à 2008 nous nous sommes attachés à étudier comment instancier et manipuler cette description formelle des spécifications.

En termes de ressources, les travaux relatifs à l'intégration ont été réalisés dans le cadre d'une thèse, d'études et de stages réalisés au COGIT, ainsi que dans le cadre d'une thèse CIFRE financée par EADS pour le PEA Envol de la DGA, et enfin dans le cadre de partenariats avec des spécialistes des ontologies et du traitement automatique du langage lors d'un projet subventionné par l'ANR GeOnto (partenaires : LRI Orsay, IRIT Toulouse, LIUPPA Pau).

PRINCIPAUX RESULTATS

Au niveau de *l'appariement de données*, de nombreux tests des outils d'appariement de réseaux mis au point dans le passé au COGIT ont été réalisés. Ces outils ont été mis à disposition à l'IGN et plus largement dans la version OpenSource de GéOxygène. Les premières utilisations mettent en avant des limites d'ergonomie pour la prise en main des outils, mais sont encourageantes sur la qualité des résultats obtenus et nous incitent à poursuivre cette diffusion des outils.

Par ailleurs, une thèse a été réalisée pour mettre au point un processus d'appariement de données relativement générique, en s'appuyant sur une gestion explicite des imperfections dans les données et les connaissances utilisées grâce à la théorie des fonctions de croyance.

Au niveau des *schémas et des ontologies*, les travaux théoriques du COGIT sur la formalisation des spécifications formelles ont été capitalisés dans GéOxygène puis concrètement mis en œuvre pour des données IGN. Tout d'abord, une taxonomie des concepts manipulés dans les spécifications IGN a été réalisée. Cette taxonomie est bilingue et contient environ 700 concepts. Ensuite, les spécifications d'une base de l'IGN ont été formalisées dans le modèle proposé et une interface de visualisation de ces spécifications formelles a été réalisée. Pour faciliter l'extraction de connaissance, des chaînes de traitement automatique du langage ont été mises en place.

La gestion de la mise à jour des données distribuées a également été étudiée dans le cadre d'une thèse CIFRE réalisée à EADS et co-encadrée au COGIT. Cette thèse a mis en avant le besoin de métadonnées de description des mises à jour pour leur intégration efficace.

De plus, l'expertise du COGIT en termes d'intégration a été exploitée dans le cadre de la directive européenne Inspire, à travers la participation à divers groupes d'experts auprès de la commission.

Enfin, de premières expériences à base de services web ont été réalisées pour préparer la mise à disposition des outils d'intégration sous la forme de services.

Pour le futur proche, les travaux sur les ontologies et l'intégration de schémas se poursuivent d'une part dans le cadre de la thèse de Nathalie Abadie débutée sur le sujet en 2008, et d'autre part dans le cadre du projet GeOnto subventionné par l'ANR et débuté en 2008 également. La participation du COGIT à la directive Inspire se poursuit aussi en 2009. Au niveau des données, les travaux sur l'appariement de données issus de la thèse d'Ana-Maria Olteanu seront repris et interfacés dans GéOxygène. Par ailleurs, une thèse d'Eric Grosso a débuté en 2008 sur la mise au point de services web de recalage de données historiques dans un référentiel récent.

Contexte

De nombreux travaux sur l'appariement de réseaux ont été réalisés au COGIT de 1995 à 2005. Certains de ces travaux sont capitalisés dans un algorithme d'appariement de réseaux à différentes échelles dans GéOxygène et déposé en OpenSource en Janvier 2008 et décrit ci-dessous.

Résultats

De nombreuses études ponctuelles d'appariement de réseaux ont été réalisées en 2006, de l'appariement de données d'un même producteur (données à différentes échelles ou de différents thèmes de l'IGN, données au 1 :50.000 et 1 :10.000 de l'IGN Belge) ou de sources plus hétérogènes (données de la Fédération Française de Randonnée Pédestre ou de l'Inventaire Forestier National avec des données de l'IGN). Plus globalement, nous avons été confrontés à une demande croissante d'expertises pour l'appariement, par exemple de la part de l'IGN France International, du Centre d'Etude sur les Réseaux les Transports et l'Urbanisme ou de la région Wallonie en Belgique. Ces demandes illustrent l'actualité de ce sujet de recherche et sont révélatrices de la diversité des utilisations de l'appariement : refonte de bases de données, migration de données utilisateur, étude de la qualité, gestion de la mise à jour, enrichissement de données, etc.

Par ailleurs, le code de l'appariement a été introduit dans la version de GéOxygène déposée en OpenSource début 2008. A cela s'est associé le nécessaire travail de conseil et de soutien à l'utilisation du code. Des premières utilisations de ce code sont à noter : l'Ordnance Survey et la région Wallonie ont réalisé des tests à partir de ce code ; l'IGN-Belge est en cours d'intégration de nos outils d'appariements dans ses lignes de production, pour appairer ses données routières au 1:10.000 et au 1:50.000 en vue de faciliter leurs mises à jour.

Par ailleurs, les algorithmes d'appariement ont également été repris et intégrés dans la chaîne de traitement mise en place par le projet « Nouvelle Carte de Base » de l'IGN (intégration dans le logiciel Clarity® de 1Spatial), pour assurer la cohérence géométrique entre les objets de divers thèmes.



Visualisation de l'appariement (ici pour le recalage des forêts de l'Inventaire Forestier National sur celles de l'IGN)

Pour en savoir plus ...

Mustière S., Devogele D. 2008. « Matching networks with different levels of detail ». *GeoInformatica*, vol.12, n.4, 12/2008, pp.435-453.

Contexte

De 2006 à 2008, Ana-Maria Olteanu a réalisé une thèse de doctorat de l'université de Paris-Est intitulée « Fusion de connaissances imparfaites pour l'appariement de données géographiques - Proposition d'une approche s'appuyant sur la théorie des fonctions de croyance », dirigée par Anne Ruas et co-encadrée par Sébastien Mustière du COGIT.

Objectifs

L'objectif de cette thèse était de proposer une méthodologie d'appariement de données géographiques générique et robuste, afin de répondre aux besoins croissant d'appariement de données de plus en plus variées et hétérogènes. Pour atteindre ces objectifs, il est nécessaire de combiner de nombreux critères d'appariement, dont certains sont très imparfaits car reposant sur la comparaison d'informations incomplètes (ex. attributs parfois non remplis), incertaines (ex. attributs parfois mal remplis), ou imprécises (ex. attributs au contenu flou ou géométrie imprécise).

Résultats

Au cours de cette thèse, un processus générique d'appariement multicritère a été défini. Il repose sur la théorie des fonctions de croyance qui présente de nombreux avantages tels que la modélisation de l'ignorance que l'on a sur un événement, et des méthodes pour la fusion de connaissances diverses, tout en modélisant le degré de désaccord entre les sources.

L'approche est en trois étapes. La première consiste à sélectionner les candidats à l'appariement de chaque objet et à les analyser individuellement. Chaque critère d'appariement (la position, les toponymes, la nature, etc.) associe un degré de croyance aux propositions suivantes : « le candidat est le bon homologue », « le candidat n'est pas le bon homologue », « ce critère ne permet pas de savoir si c'est ou non un bon candidat ». La deuxième étape consiste à fusionner les croyances par candidat et la troisième étape consiste à fusionner les croyances sur tous les candidats afin d'avoir une vue d'ensemble et de prendre une décision.

La méthode a été mise en œuvre d'une part pour l'appariement des points remarquables du relief de deux bases de données de l'IGN à différentes échelles, et d'autre part pour l'appariement de données routières à différentes échelles et issues de différents producteurs (cf. image ci-jointe). Dans ces tests, des critères de localisation, de nature, de toponymie, ou liés à la topologie ont été utilisés. On peut constater que l'appariement fonctionne malgré la différence de niveau de détail des deux bases.



Cette thèse a montré l'intérêt de modéliser le degré d'incertitude des connaissances utilisées pour appairer. Elle a surtout montré que l'utilisation conjointe de nombreux critères d'appariement était possible et très efficace. Les outils d'appariement de données définis seront exploités dans les thèses débutées en 2008, de Nathalie Abadie sur l'appariement de schémas, et d'Eric Grosso sur le recalage de données historiques sur un référentiel récent.

Pour en savoir plus ...

Olteanu-Raimond A.-M., 2008. « Fusion de connaissances imparfaites pour l'appariement de données géographiques - Proposition d'une approche s'appuyant sur la théorie des fonctions de croyance ». Thèse de Doctorat de l'Université Paris Est, soutenue le 24 octobre 2008.

Olteanu-Raimond A.-M., Mustière S. 2008. « Data matching - a matter of belief ». Proceedings of the International Symposium on Spatial Data Handling SDH'2008, Montpellier, pp. 501-519.

Contexte

Pour résoudre divers problèmes liés à l'hétérogénéité sémantique des données, il est utile de disposer d'ontologies qui permettent de spécifier sans ambiguïté la sémantique de systèmes terminologiques donnés, en particulier ceux utilisés pour la construction de bases de données géographiques ou métiers.

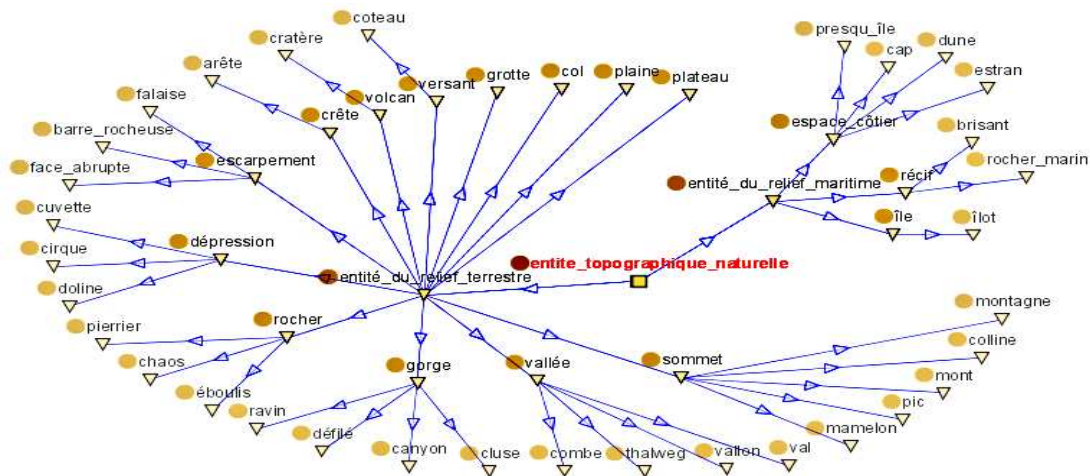
Objectifs

Notre objectif est de réaliser une ou plusieurs ontologies reflétant un point de vue topographique. Plus précisément, nous voulons dans un premier temps réaliser une taxonomie des termes utilisés à l'IGN pour décrire le contenu des bases de données, pour ensuite l'étendre à d'autres points de vue.

Résultats

Pour réaliser cette taxonomie, nous nous sommes appuyés sur les spécifications des bases de données de l'IGN, qui ont l'avantage de contenir de nombreux mots propres au vocabulaire de la géographie. La terminologie utilisée dans les spécifications est en effet beaucoup plus riche que celle utilisée dans les schémas, puisque chaque classe y est décrite en fonction des entités du monde réel qu'elle représente. Par exemple, la description de la classe « tronçon hydrographique » fait appel aux concepts de rivière, fleuve, ru, ruisseau, fossé, aqueduc.

Des taxonomies ont été réalisées à partir des spécifications de deux bases de données, en s'appuyant sur des outils de traitement automatique du langage dans le cadre d'un stage de F. Laurens en 2006. Ces taxonomies ont ensuite été fusionnées, grâce à des outils d'alignement d'ontologies et à une réorganisation interactive. La taxonomie ainsi créée contient 700 concepts, associés chacun à un ou plusieurs termes (voir un extrait concernant le relief ci-dessous). Ces termes ont été traduits en anglais pour obtenir une taxonomie bilingue qui peut maintenant servir de base à différents travaux pour son exploitation et son enrichissement.



Cette première chaîne de traitements des spécifications textuelles a été réalisée à partir d'outils ad-hoc adaptés à la forme des spécifications traitées. Une nouvelle chaîne de traitements, plus générique, ainsi qu'un processus d'alignement automatique et d'enrichissement d'ontologies, sont en cours de réalisation dans le cadre du projet GeOnto. Pour cela, nous avons en particulier décidé de définir un format d'entrée des spécifications s'appuyant sur la norme ISO 19131 (Geographic Information – Data Product Specification), norme utilisée en particulier pour les spécifications définies dans le cadre de la directive INSPIRE.

Pour en savoir plus ...

Abadie N. et Mustière S. 2008. « Constitution d'une taxonomie géographique à partir des spécifications de bases de données ». Actes de SAGEO'2008, Montpellier, 26 juin 2008.

Site Web du projet GeOnto : <http://geonto.lri.fr>

Contexte

Dans le cas des bases de données géographiques, les informations décrivant les données se situent à plusieurs niveaux. Même si, en première approche, on se réfère aux schémas conceptuels des bases de données pour en comprendre le contenu, la majeure partie de la sémantique des bases de données géographiques se trouve dans les spécifications mêmes de ces bases. Celles-ci, en effet, décrivent la sémantique du contenu d'une base à travers le processus d'acquisition des données.

Diverses applications nécessitent de comprendre cette sémantique. C'est le cas notamment de l'aide à la description du contenu des bases pour les utilisateurs et de l'intégration de données hétérogènes. Pour que ces applications puissent être automatisées ou au moins assistées, il nous semble nécessaire de formaliser ces spécifications. Un modèle pour cela a été défini au COGIT dans la thèse de Nils Gesbert soutenue en 2005.

Objectifs

Un objectif de nos travaux a été en premier lieu de mettre en œuvre ces travaux théoriques de formalisation dans la plateforme GéOxygène. Ensuite il a fallu instancier ce modèle formel pour des bases de données IGN, et plus globalement étudier et concevoir des outils facilitant cette instanciation.

Résultats

Les spécifications de la BDTOPPO de l'IGN ont été formalisées en grande partie, avec une forte part d'automatisation grâce à des outils de traitement automatique du langage lors d'un stage réalisé par V. Picard en 2007.

Par ailleurs, une première interface de consultation de spécifications formelles a été réalisée dans un stage de T. Horel réalisé en 2008 (image ci-contre).

Cette interface permet de visualiser à la fois une ontologie, un schéma de base de données, des spécifications textuelles brutes, et leur formalisation dans un langage s'appuyant sur XML.

Ces spécifications vont maintenant être utilisées dans le cadre de la thèse de Nathalie Abadie, débutée en 2008, sur l'appariement de schémas. Le premier cas d'utilisation est l'intégration de données utilisateur quelconques, peu décrites, avec des données de référence richement décrites par ces spécifications formelles.

Pour en savoir plus ...

Abadie N. et Mustière S. 2008. « Constitution d'une taxonomie géographique à partir des spécifications de bases de données ». Actes de SAGEO'2008, Montpellier, 26 juin 2008.

PARTICIPATION A LA DIRECTIVE EUROPEENNE INSPIRE

Anne Ruas et Sébastien Mustière

Contexte

INSPIRE (INfrastructure for SPatial InfoRmation in Europe) est une directive européenne qui vise à fixer les règles générales pour établir une infrastructure d'information géographique dans la Communauté Européenne aux fins des politiques environnementales communautaires. La directive a été adoptée le 14 mars 2007. Pour la définition de spécifications de données européennes harmonisées entre autres, des règles de mises en œuvre sont en cours de définition de 2008 à 2012.



Objectifs et résultats

Pour nos recherches, la directive constitue un exemple riche d'intégration de données transfrontalières. De plus, la participation à cette directive est importante pour l'IGN qui y sera concrètement confronté. Pour toutes ces raisons, et en tant que spécialiste de l'intégration de données géographiques, le COGIT participe auprès de la commission européenne à la définition des règles de mise en œuvre de la directive. Anne Ruas a été jusqu'en 2008 membre de l'équipe de rédaction de la méthodologie de définition des spécifications harmonisées européennes. Sébastien Mustière est, quant à lui, rédacteur en 2008 et 2009 de ces spécifications sur le thème de la toponymie.

Pour en savoir plus ...

Site web de la commission européenne sur Inspire : <http://www.ec-gis.org/inspire/>

GESTION DE MISES A JOUR CONCURRENTES

Christelle Pierkot

Contexte

De 2004 à 2008, Christelle Pierkot a réalisé une thèse à l'université Paul Sabatier de Toulouse sur la gestion des mises à jour concurrentes dans les bases de données géographiques. La thèse financée par EADS dans le cadre du volet Cohérence du PEA ENVOL commandité par la DGA, était dirigée par Abdelkader Hameulain (IRIT) et Anne Ruas (COGIT) et co-encadrée par Sébastien Mustière (COGIT). La thèse a été réalisée à EADS Toulouse et au COGIT durant 6 mois.

Objectifs

L'objectif de la thèse était d'améliorer la cohérence des mises à jour des données géographiques dans un contexte où ces données sont réparties et potentiellement mises à jour par divers acteurs de manières concurrentes. Plus concrètement, le contexte étudié était celui d'une mission militaire, où les données sont distribuées entre équipes de terrain et quartiers généraux, et où les mises à jour proviennent de missions terrain, de saisie par télédétection ou de sources externes.

Résultats

Un processus d'intégration de mises à jour concurrentes a été défini. L'approche s'appuie sur la définition d'un modèle de métadonnées des mises à jour, qui étend le modèle standard de métadonnées de l'ISO19115. Ce modèle permet de représenter les rôles et caractéristiques des données, acteurs et évolutions. En termes de processus, l'approche proposée s'appuie sur les travaux relatifs à la réconciliation dans les bases de données. Elle repose sur l'alternance de phases de gestion de la concurrence et de phases d'intégration des mises à jour réconciliées. L'approche a été formalisée dans le langage ECA (Événement, Condition, Action) qui est principalement dédié à la description des bases de données actives.

Pour en savoir plus ...

Pierkot C., 2008. « Gestion de la Mise à Jour de Données Géographiques Répliquées ». Thèse de Doctorat de l'université Paul Sabatier Toulouse III, soutenue le 2 juillet 2008

Rédigé par Cécile Duchêne, responsable de l'action de recherche

CONTEXTE ET OBJECTIFS

La généralisation de données géographiques est le processus qui consiste à simplifier l'information contenue dans une base de données géographique pour répondre à un besoin donné. Ce besoin peut être par exemple l'analyse d'un phénomène géographique à un niveau de détail moindre que celui de la base de données initiale, ou encore l'affichage des données de manière lisible à une échelle plus petite que l'échelle de référence de la base initiale (généralisation à but cartographique). La généralisation s'apparente à un résumé de texte. Elle diminue le niveau de détail des données d'origine, tout en gardant le sens initial des données. De plus, elle met en exergue les éléments importants au regard du besoin auquel elle répond.

Les recherches menées au laboratoire COGIT sur ce sujet ont pour but de répondre aux besoins suivants présentés dans le Schéma Directeur de la Recherche et des Développement 2006-2010 de l'IGN :

- la valorisation de données, interne à l'IGN : gérer de façon optimale des bases de données à différentes échelles (généralisation non cartographique), optimiser la fabrication de cartes à partir de bases de données,
- la valorisation externe « grand public » : fabriquer et diffuser des produits cartographiques « à la carte »,
- la valorisation externe vers les professionnels : construire des référentiels de différents niveaux de détail en fonction des besoins, notamment dans le contexte européen de la directive INSPIRE.

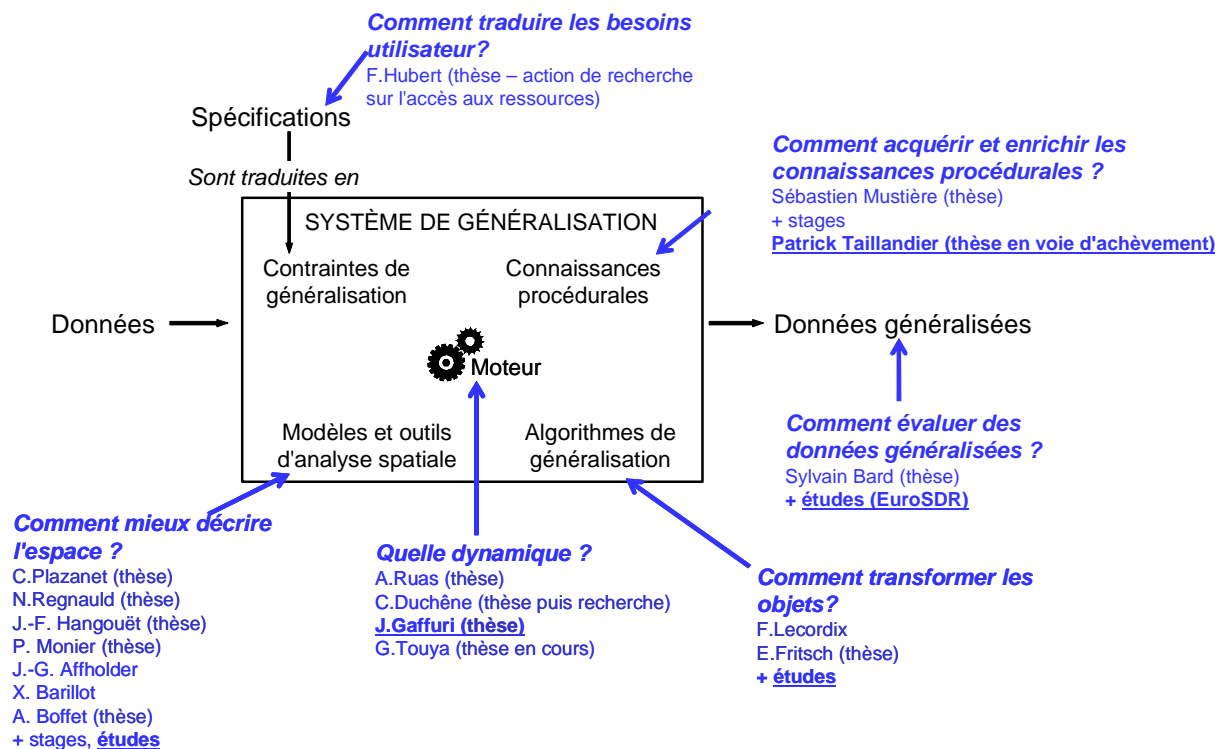
Pour répondre à ces besoins, il est nécessaire de disposer de processus de généralisation automatique suffisamment fiables, souples et paramétrables sachant répondre à différents besoins de généralisation.

Des recherches en généralisation sont menées au laboratoire COGIT depuis 1991, en lien avec la communauté internationale de recherche au sein de laquelle le COGIT joue un rôle actif. Sur la période 1991-2006, les travaux au COGIT ont d'abord porté sur la mise au point d'algorithmes de généralisation et sur l'enrichissement des données à l'aide d'outils d'analyse spatiale, de manière à mieux guider la généralisation (détection du besoin de généralisation, préservation des caractéristiques principales des données). Nous avons également travaillé sur l'évaluation des données généralisées. Sur ces trois thèmes confondus, sept thèses ont été soutenues entre 1996 et 2004 (cf. figure page suivante).

Depuis 1998, l'essentiel de notre effort de recherche s'est porté sur l'orchestration du processus de généralisation, que nous séparons sur la figure suivante en deux volets : (1) la dynamique du processus, avec la proposition de modèles aussi génériques que possible permettant d'enchaîner les algorithmes de généralisation, guidés par des connaissances (heuristiques) (2) l'acquisition et la révision des connaissances destinées à guider ces modèles. Trois modèles de généralisation complémentaires ont ainsi été proposés :

- AGENT en 1999 pour les zones denses type zones urbaines (thèse d'Anne Ruas, puis projet européen AGENT),
- CartACom en 2004 pour les zones peu denses et hétérogènes comme les zones rurales (thèse de Cécile Duchêne), et
- GAEL, pour les thèmes de fond de carte (thèse de Julien Gaffuri, qui était en cours en 2006).

Un modèle utilisant de l'apprentissage automatique pour l'acquisition des connaissances de guidage à partir d'exemples recueillis auprès d'experts a été proposé (thèse de Sébastien Mustière, 2001). Début 2006, une thèse commençait au laboratoire sur la révision automatique des connaissances contenues dans un système de généralisation, par analyse des traces (thèse de Patrick Taillandier).



Situation des travaux effectués en généralisation depuis 1991. En caractères gras les travaux sur la période 2006-2008 qui sont détaillés dans la suite du document.

Sur la période 2006-2008, nous avons essentiellement focalisé nos recherches sur les trois aspects suivants :

- Outils d'analyse spatiale et algorithmes de généralisation : la sélection de réseaux. Il s'agit d'améliorer et de rendre plus génériques les algorithmes de sélection de réseaux existants, (1) pour faire de la généralisation à but non cartographique d'une base de données vers une autre avec certaines contraintes, ou (2) en préalable à la généralisation cartographique avec saut d'échelle important. Cet objectif avait été identifié lors d'une étude technique menée en 2005 au laboratoire,
- Dynamique du système de généralisation : un nouveau modèle pour la prise en compte des thèmes de fond de carte. Ces thèmes, tels que le relief et l'occupation du sol, sont souvent négligés dans les recherches en généralisation automatique. Il s'agit de savoir tenir compte de ces thèmes lors de la généralisation des thèmes « principaux », de manière à préserver les relations inter-thèmes. Pour cela, un modèle particulier est nécessaire, qui doit être coordonnable avec les modèles existants pour les thèmes principaux,
- Acquisition et enrichissement des connaissances procédurales : la révision automatique des connaissances procédurales contenues dans un système de généralisation. Il s'agit de pouvoir faire évoluer un système de généralisation à base de connaissances en révisant automatiquement les connaissances qu'il contient.

APPROCHES

Approche générale du laboratoire COGIT en généralisation

Les recherches en généralisation menées au laboratoire COGIT font appel à deux disciplines scientifiques : d'une part la géomatique, notamment la géométrie algorithmique, d'autre part l'Intelligence Artificielle, notamment les systèmes multi-agents et l'apprentissage artificiel.

L'approche de la généralisation adoptée au laboratoire COGIT est depuis les années 1990 une approche locale, pas à pas, à base de connaissances, selon les modèles théoriques proposés en 1988 en Suisse par Brassel et Weibel d'une part, aux Etats-Unis par McMaster et Shea d'autre part. Le principe est de caractériser les données pour détecter les endroits où il y a un besoin de généralisation, et

d'appliquer localement des algorithmes adaptés en fonction des conflits détectés et des caractéristiques des données.

Le modèle AGENT, proposé dans la thèse d'Anne Ruas (1999) puis ré-implémenté et enrichi lors du projet européen AGENT (achevé en 2000), reprend les principes ci-dessus et les traduit en un modèle opérationnel utilisant le paradigme des systèmes multi-agents. Les objets géographiques y sont modélisés comme des agents, c'est-à-dire des entités autonomes mues par un but à atteindre. Ces agents géographiques ont pour but de satisfaire un certain nombre de contraintes (par exemple, un bâtiment désire avoir une taille minimale), qui traduisent les spécifications du résultat souhaité de la généralisation. Ces agents géographiques vectoriels ont des capacités d'introspection leur permettant d'évaluer le degré de satisfaction de leurs contraintes, et réalisent des cycles 1/évaluation 2/choix d'un algorithme de généralisation 3/auto-application de l'algorithme 4/réévaluation 5/validation ou annulation/ etc. De plus, dans le modèle AGENT, plusieurs niveaux d'agents sont considérés : les agents micro représentent les objets géographiques élémentaires, et les agents meso représentent des groupes organisés d'objets géographiques. Les interactions entre agents sont hiérarchiques (du groupe vers ses composants et vice versa). Ce modèle fonctionne bien pour la généralisation de données bien hiérarchisées, comme l'espace urbain aux moyennes échelles ou encore les routes de montagne qui sont généralisées après une décomposition en portions homogènes en termes de conflits.

Le modèle CartACom, proposé dans la thèse de Cécile Duchêne (2004), reprend les principes du modèle AGENT (objets géographiques modélisés comme des agents, guidés par des contraintes). Il propose une modélisation des contraintes relationnelles, partagées par deux agents (contrainte de non superposition, contrainte de préservation du parallélisme), et introduit des interactions transversales entre agents pour la prise en compte de ces contraintes. Ce modèle fonctionne bien pour la généralisation de données hétérogènes en zone peu dense, comme l'espace rural aux moyennes échelles.

Dans les modèles AGENT et CartACom, le cycle d'action des agents géographiques est guidé par de nombreuses connaissances qui traduisent des heuristiques de recherche d'une solution : choix de l'agent à activer, choix de la ou des contraintes à traiter en priorité, choix de l'action à s'appliquer, validation ou abandon de la transformation effectuée, etc. La qualité de ces connaissances a un impact direct sur la qualité du résultat obtenu (efficacité) et sur la vitesse d'obtention du résultat (efficacité), beaucoup d'algorithmes de généralisation étant gourmands en temps de calcul. Or la mise au point de ces connaissances « à la main » est longue et fastidieuse, et elles doivent de plus être révisées lorsque le système évolue, par exemple par ajout d'un nouvel algorithme de généralisation. La thèse de Sébastien Mustière (2001) a proposé un modèle pour l'acquisition automatique de connaissances de guidage à partir d'exemples saisis par un expert, mais cela pose le problème du recueil d'exemples. Pour dépasser ce problème, nous étudions deux approches complémentaires. La première consiste à développer un module d'aide à la saisie et à l'analyse de connaissances d'experts (un stage réalisé en 2004, d'autres prévus à court terme). La seconde approche consiste à constituer des bases d'exemples automatiquement, par analyse des traces d'exécution du système de généralisation : il s'agit alors de réviser les connaissances contenues dans le système à partir de l'expérience. Cette seconde approche a été étudiée durant la période 2006-2008 sous la forme d'une thèse de doctorat (Patrick Taillandier).

Démarche sur la période 2006-2008

Le travail sur la sélection de réseaux a constitué la tâche principale de Guillaume Touya pendant les années 2006-2007. Outre le codage sur notre plate-forme de recherche Clarity d'une structure de graphe topologique souple et manipulable, ce travail a consisté à reprendre, adapter et compléter diverses méthodes de détection de structures caractéristiques des réseaux issues de la littérature récente, ainsi que des méthodes de sélection proprement dites. L'ensemble de briques ainsi constitué a ensuite été organisé en un processus de sélection complet.

En ce qui concerne la dynamique du processus de généralisation, le principal point étudié sur la période 2006-2008 a été la prise en compte de thèmes de fond de carte dans un processus de généralisation. Cela a fait l'objet de la thèse de Julien Gaffuri, qui s'est déroulée d'octobre 2004 à

juillet 2008. Cette thèse a porté sur la mise au point d'un modèle permettant aux thèmes de fond de carte de type « champ » définis en tout point de l'espace, comme le relief ou l'occupation du sol, d'avoir un comportement élastique leur permettant (1) d'absorber, lorsque c'est possible, les déformations induites par les thèmes de premier plan (routes, bâtis, hydro, etc.), (2) de contraindre en retour ces thèmes de premier plan, lorsque la déformation engendrée est jugée trop importante. En termes de résultats nous avons produit : 1- un modèle multi-agents GAEL pour la déformation contrainte d'objets champs, 2- la mesure de la quantité de déformation d'un champ et les mécanismes de rétroaction possibles sur les objets principaux, et enfin 3- le couplage du modèle de déformation proposé avec les modèles de généralisation dédiés aux thèmes principaux déjà développés au laboratoire COGIT (AGENT et CartACom).

Le couplage entre les trois modèles (AGENT, CartACom et GAEL) a été étudié de manière plus large par Julien Gaffuri et Cécile Duchêne : ont été analysées, les similitudes et différences entre ces trois modèles (dont les principes sont assez proches), et la faisabilité d'un socle commun permettant de les utiliser conjointement. De manière complémentaire, la thèse de Guillaume Touya, commencée fin 2007, porte sur l'orchestration d'un processus de généralisation faisant appel à des processus basés sur des principes différents, déclenchés séparément sur différentes parties (espace ou thème) d'un jeu de données. La première année de ce travail a permis de cerner les problématiques sous-jacentes : expression des contraintes dans un formalisme interopérable avec différents modèles de généralisation, partitionnement des données, gestion des raccords aux frontières, orchestration proprement dite, harmonisation des résultats et sensibilité des différents modèles.

La thématique de l'acquisition et de l'enrichissement des connaissances procédurales a été étudiée dans le cadre de la thèse de Patrick Taillandier, qui s'est déroulée d'octobre 2005 à décembre 2008. La problématique étudiée dans cette thèse est la révision automatique des connaissances procédurales contenues dans un système de généralisation basé sur l'exploration d'arbres d'états, comme c'est le cas du système AGENT : chaque objet géographique qui se généralise construit, au cours de son cycle d'action, un arbre d'états dans lequel un état correspond à une solution possible, et une transition correspond à une opération. Les arbres d'états, construits par les objets géographiques au cours de leur généralisation, sont une source pour apprendre les bonnes connaissances de guidage (car elles mènent à un résultat satisfaisant), ou au contraire défectueuses. Les points étudiés dans ce travail de thèse sont la constitution d'un jeu d'exemples représentatif, la modélisation des connaissances sous une forme manipulable par une machine, le diagnostic en ligne des connaissances préalable au déclenchement d'un processus de révision, et la révision proprement dite des connaissances à partir de l'expérience que constituent les exemples préalablement généralisés.

En dehors des trois thèmes de recherche décrits ci-dessus, deux autres travaux ont été menés.

- Le premier se situe dans le cadre d'un projet mené sous l'égide d'EuroSDR et coordonné par Jantien Stöter, de l'ITC (Entschede, Pays-Bas). Il s'agit de tester différents logiciels de généralisation disponibles sur le marché. Nous avons pris une part active à ce projet, tant dans la définition du protocole de test que dans la réalisation du test proprement dite et dans l'évaluation des résultats. La phase d'évaluation est actuellement en cours, et le projet devrait s'achever début 2009.
- Le deuxième travail est une étude technique réalisée à la demande du Service des Bases de données Vecteur de l'IGN, dans le cadre de la mise en place d'une collaboration entre l'IGN et l'Inventaire Forestier National (IFN) pour la saisie et l'exploitation d'un « fond vert commun » (données de végétation). Cette étude a porté sur les traitements de généralisation possibles et souhaitables pour améliorer le « fond vert commun » afin de faciliter son exploitation à l'IGN et à l'IFN.

PRINCIPAUX RESULTATS

Résultats de recherche

1- Des processus et algorithmes de sélection de réseaux

Le travail de Guillaume Touya sur la sélection de réseaux a abouti à l'implémentation de processus paramétrables de sélection des réseaux routier et hydrographique, basés sur l'intégration de nombreux travaux issus de la littérature. Ces algorithmes ont été testés pour le passage de données topographiques grande échelle (BD TOPO®) à plusieurs des données moyennes échelles avec des granularités variables, et les résultats obtenus sont jugés satisfaisants.

2- Un modèle de déformation de l'espace pour le maintien des relations pendant le processus de généralisation

La thèse de Julien Gaffuri, soutenue en juillet 2008, a défini un modèle permettant de répercuter sur les thèmes de fond de carte de type « champ » (comme le relief), les effets de la généralisation des thèmes principaux. Ce modèle, baptisé GAEL pour « Généralisation à base d'Agents ELastiques », inclut une modélisation des contraintes intra-champ (préservation de la forme notamment) et des contraintes entre champs et objets vectoriels, et des modèles dynamiques pour la déformation sous contrainte d'un objet de type *champ*, et pour les interactions entre objets vectoriels et objets *champs* lors du processus de généralisation. Il a été validé par des tests sur des données topographiques pour les applications suivantes : (1) le maintien de la relation de position relative entre un bâtiment et le relief lors de la généralisation contextuelle du bâtiment, (2) le maintien, voire l'amélioration, de la cohérence entre l'hydrographie et le relief, et (3) le maintien des positions relatives entre les thèmes principaux et l'occupation du sol.

3- Un premier couplage de modèles de généralisation

Un travail sur le couplage des modèles de généralisation a été entrepris sur la fin de cette période 2006-2008, d'une part par Cécile Duchêne et Julien Gaffuri pour ce qui est du rapprochement des modèles du COGIT, d'autre part par Guillaume Touya dans le cadre de sa thèse commencée fin 2007, pour ce qui est de l'orchestration d'un processus de généralisation faisant appel à des modèles différents. A ce stade, nous estimons qu'une bonne analyse des problématiques a été réalisée sur ces deux aspects. Ces travaux seront poursuivis activement durant les années à venir.

4- Un modèle de révision des connaissances de généralisation pour assurer la pérennité des modèles

La thèse de Patrick Taillandier, dont la soutenance est prévue en décembre 2008, a défini un modèle pour la révision automatique des connaissances de guidage contenues dans un système de généralisation basé sur l'exploration d'arbres d'états. Au cours de cette thèse un processus complet de révision de connaissances pour l'automatisation du processus de généralisation à bases d'agents a été conçu et implémenté : un modèle de stockage des connaissances, une méthode de choix automatique d'exemples représentatifs, un module de traçage, une méthode de révision proprement dite et enfin une méthode de diagnostic, permettant de juger en ligne s'il est nécessaire de réviser les connaissances. Ce travail nous permet, non seulement de disposer effectivement d'un module de révision des connaissances dans le modèle de généralisation AGENT comme escompté, mais aussi d'avoir à disposition diverses bibliothèques d'apprentissage automatique et l'expertise pour les interfacer avec notre plate-forme de recherche Clarity.

5- Evaluation de données et de processus

Le travail réalisé dans le cadre du projet de tests de logiciels de généralisation mené sous l'égide d'EuroSDR nous a permis notamment de mettre au point et de tester des principes méthodologiques sur l'évaluation de données généralisées. Cela a été aussi l'opportunité de réutiliser l'expertise issue de travaux antérieurs menés, entre autres, au laboratoire COGIT : précédent test coordonné par Anne Ruas sous l'égide de l'OEEPE (1997-2001) pour ce qui est du protocole de test, projet européen AGENT et thèse de Sylvain Bard (2004) respectivement pour l'évaluation visuelle de résultats par des experts et pour l'évaluation automatique.

Plate-forme de recherche

La période 2006-2008 a vu notre plate-forme de recherche évoluer, passant du SIG LAMPS2 au logiciel de généralisation Clarity. Les deux logiciels sont commercialisés par la société anglaise 1Spatial (anciennement Laser-Scan). LAMPS2 est la plate-forme sur laquelle a été développé le prototype de généralisation mis au point pendant le projet européen AGENT (1997-2000). La plupart des travaux réalisés en généralisation au COGIT antérieurement ont été portés sur LAMPS2 pendant le projet AGENT (en particulier, de nombreux algorithmes de généralisation et le modèle AGENT issu de la thèse d'Anne Ruas). LAMPS2 est alors devenu notre plate-forme de recherche, et les travaux en généralisation ont été développés sur cette forme jusqu'en 2005.

A l'issue du projet AGENT, 1Spatial a décidé de remettre à plat le prototype de généralisation développé sur LAMPS2 durant le projet. Le résultat de cette remise à plat est le logiciel de généralisation **Clarity**, qui inclut une nouvelle implémentation du modèle AGENT et des algorithmes de généralisation présents sur LAMPS2 en fin de projet AGENT.

Clarity fournit une meilleure interface de paramétrage du modèle de généralisation AGENT, et un meilleur environnement de développement en Java. Une version jugée suffisamment stable de Clarity a été disponible en 2005. Nous avons alors décidé d'adopter Clarity comme nouvelle plate-forme de recherche. La plupart des travaux que nous avons développés sur LAMPS2 ont pu être portés sur Clarity par encapsulation (c'est-à-dire sans recodage lourd). Le fait que le langage de programmation de Clarity soit Java, comme dans l'autre plate-forme de recherche du laboratoire GeOxygène, nous permet désormais d'aller vers une mutualisation croissante des outils développés sur ces deux plates-formes de recherche.

Passage de travaux en production

Le travail de thèse de Julien Gaffuri a été partiellement repris par le projet Nouvelle Carte de Base de l'IGN, qui met au point la future chaîne de production des cartes au 1 : 25 000 et au 1 : 50 000 à partir de la BDTopo®. Les applications sont la mise en cohérence de l'hydrographie et du relief et l'interfaçage des processus de généralisation du routier et du bâti.

Le travail réalisé par Guillaume Touya sur la sélection de réseaux est actuellement repris pour la dérivation de la composante française des produits européens EuroRegionalMap et EuroGlobalMap à partir de la BD CARTO®. Ce travail sera à nouveau utilisé dans le cadre du projet européen ESDIN auquel participe l'IGN, portant sur la constitution ou l'amélioration des référentiels européens à petite et moyenne échelles dans le cadre de la directive européenne INSPIRE.

Enfin, le travail réalisé sur la généralisation du « fond vert commun » co-produit par l'IGN et l'IFN a été jugé suffisamment intéressant pour être repris en production à l'IGN. L'industrialisation du code développé au COGIT dans le cadre de cette étude technique a été réalisée par le Projet Nouvelle Carte de Base. Une partie du processus mis au point dans l'étude a été adoptée en vue d'améliorer la couche végétation du RGE : il s'agit de différencier les forêts des haies et bosquets, de partitionner les forêts par les réseaux et de mettre en cohérence les limites de forêts avec les réseaux (ce dernier point faisant appel à des outils d'appariement du COGIT). La mise en œuvre de ces propositions a abouti à la modification des spécifications de la couche végétation du RGE, et à la mise en œuvre des algorithmes fournis au Service des Bases de données Vecteur de l'IGN. Le reste du processus proposé, à savoir la squelettisation des haies, a été utilisé par le projet Carte de Base de l'IGN pour la dérivation de la carte de base à partir du RGE.

Contexte

Suite à l'étude sur l'amélioration de la BDCarto à partir du RGE réalisée par le COGIT en 2004 et 2005, il est apparu nécessaire et faisable de compléter les processus de sélection de réseau développés durant l'étude pour aller vers plus de généralité. La nécessité d'explicitier les structures caractéristiques des réseaux a notamment été soulignée, les critères utilisés pour spécifier le résultat attendu de la sélection faisant souvent référence à ces structures caractéristiques.

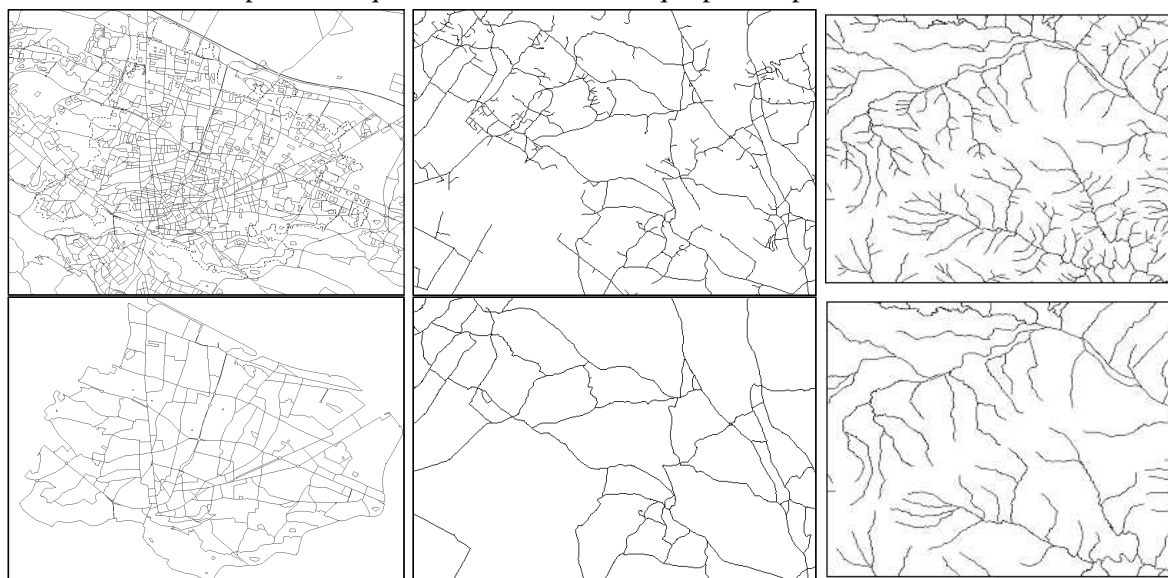
Objectif

L'objectif de ce travail est de concevoir un processus paramétrable de sélection de réseaux utilisant les structures géographiques implicites dans une base de données topographiques. Ce processus doit être appliqué aux réseaux routier et hydrographique.

Résultats

Guillaume Touya a développé un processus générique en reprenant des travaux de la littérature, en les enrichissant, en complétant les manques et surtout en assemblant ces briques en un processus cohérent. Il a ainsi proposé un processus en trois étapes :

1. L'enrichissement de la base de données initiale afin de créer de nouveaux objets par des méthodes d'analyse spatiale pour explicitier le contexte géographique et identifier des structures caractéristiques (un échangeur routier ou un delta par exemple).
2. La sélection basée sur les enrichissements. Pour le réseau routier, deux algorithmes différents sont utilisés pour la sélection en zone rurale et urbaine (voir figure ci-dessous), et un contrôle de cohérence à la frontière est opéré ensuite.
3. Le traitement adapté de chaque structure caractéristique pour la préserver durant la sélection.



(a) Zone urbaine

(b) Zone rurale

(c) Hydrographie

Résultats de sélection obtenus sur un réseau routier (a) et (b) et sur un réseau hydrographique (c)

Pour en savoir plus ...

Touya G. 2007. "A Road Network Selection Process Based on Data Enrichment and Structure Detection". In proceedings of 10th ICA Workshop on Generalisation and Multiple Representation. Moscow (Russia). 2007

Touya G. 2007. "River Network Selection based on Structure and Pattern Recognition". In proceedings of ICC2007, the International Cartography Association Conference, Moscow (Russia), 2007.

Touya G. 2008. "Enrichissement de données par analyse spatiale pour la généralisation des réseaux". Colloque International de Géomatique et d'Analyse Spatiale SAGEO'08, Montpellier, France.

Contexte

Thèse de doctorat réalisée d'octobre 2005 à décembre 2008. Université Paris-Est. Directeur de thèse : Alexis Drogoul (IRD), Co-encadrement : Cécile Duchêne (COGIT).

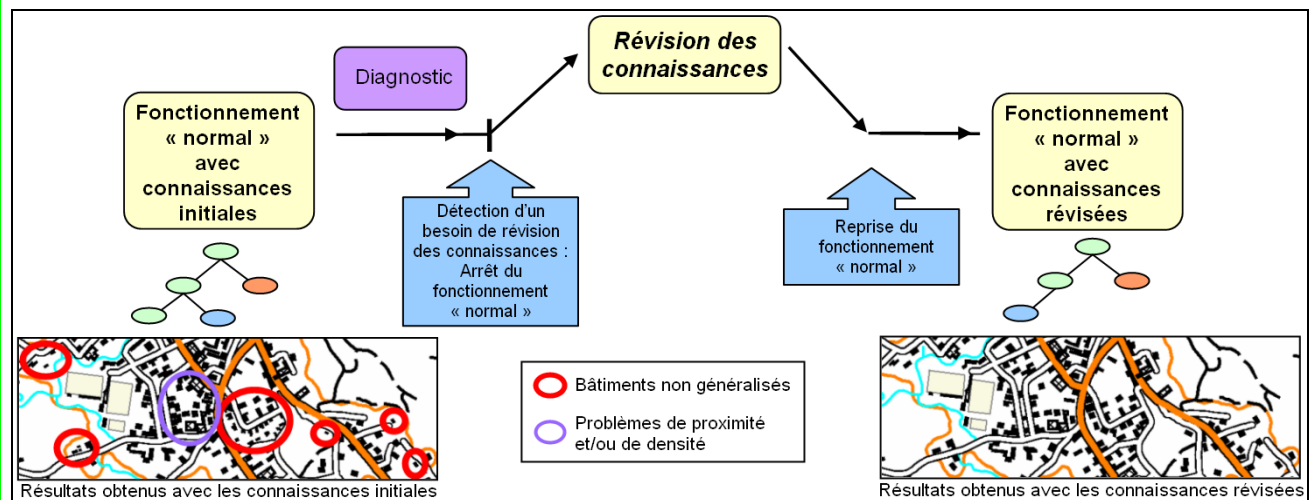
Ce travail de thèse concerne l'acquisition et l'enrichissement des connaissances procédurales contenues dans les systèmes de généralisation. Les connaissances procédurales sont particulièrement importantes pour les systèmes fonctionnant par exploration d'arbres d'états comme c'est le cas des systèmes de généralisation basés sur le modèle AGENT. La mise au point de ces connaissances, qui est nécessaire à chaque évolution du système, s'avère fastidieuse. Des méthodes ont été proposées pour acquérir automatiquement de telles connaissances (thèse de S. Mustière en 2001), mais celles-ci nécessitent un trop long travail de collecte d'exemples.

Objectifs

L'objectif de ce travail est de proposer une approche permettant la révision automatique des connaissances procédurales contenues dans les systèmes fonctionnant par exploration informée d'arbres d'états et d'en proposer une mise en œuvre pour le modèle AGENT.

Résultats

Patrick Taillandier a proposé une approche générale de révision par l'expérience des connaissances guidant l'exploration d'arbres d'états. L'approche proposée, qui fonctionne hors ligne, se base sur le traçage du système pendant qu'il généralise un grand nombre d'objets géographiques et sur l'analyse de ces traces pour modifier en conséquence les connaissances du système. Il a également proposé une approche de diagnostic en ligne de la qualité des connaissances permettant de déterminer quand déclencher le processus de révision hors ligne des connaissances. Ces deux approches ont été mises en œuvre et expérimentées dans le cadre du modèle AGENT. Les résultats obtenus ont montré que les connaissances révisées permettaient au système de mieux généraliser, ou de généraliser plus vite sans baisse de qualité.



Approche générale de révision des connaissances dans la thèse de Patrick Taillandier

Pour en savoir plus ...

Taillandier P. 2008. « Révision des connaissances guidant l'exploration informée d'arbres d'états. Application au contexte de la généralisation de données géographiques ». Thèse de doctorat, Université Paris-Est, 2008.

Taillandier P., Duchêne C. & Drogoul A. 2008. « Knowledge revision in systems based on an informed tree search strategy: application to cartographic generalisation », International Conference on Soft Computing as Transdisciplinary Science and Technology (CSTST), 2008.

Contexte

Thèse de doctorat effectuée de décembre 2004 à juillet 2008. Université Paris-Est. Directrice de thèse : Anne Ruas (COGIT), co-encadrement : Cécile Duchêne (COGIT).

Les modèles de généralisation automatique à base d'agents géographiques mis au point au laboratoire COGIT ont été appliqués avec succès à la généralisation des thèmes composés d'objets géographiques discrets (routier, bâti, etc.). Les thèmes « champ » tels que le relief et l'occupation du sol, qui ont pour particularité d'être définis en tout point de l'espace, ne sont pas pris en compte lors du processus de généralisation des objets discrets. En conséquence, les nombreuses relations entre objets et champs peuvent être détruites à l'issue de la généralisation des objets. Par exemple, en l'absence de modèle, un tronçon de cours d'eau en conflit de superposition avec un autre objet peut être déplacé et ainsi sortir du thalweg dans laquelle il s'écoulait initialement.

Objectif

L'objectif de ce travail de thèse est de proposer un modèle qui permette la prise en compte des thèmes champ dans un processus de généralisation automatique de manière à préserver leurs relations avec les objets discrets.

Résultats

Le nouveau modèle de généralisation baptisé GAEL (Généralisation à base d'Agents ELastiques) permet des interactions entre objets géographiques et champs pour préserver leurs relations. Par l'utilisation d'une méthode de déformation à base d'agents, les champs ont la capacité de se déformer localement sous les objets généralisés pour préserver les relations qu'ils partagent avec eux. Par exemple, le relief peut se déformer sous les cours d'eau pour préserver la relation d'écoulement qu'ils partagent. De même, le relief peut se déformer sous l'action de bâtiments pour permettre de préserver leur position relative (cf. figure ci-dessous).

Ces déformations élastiques sont faites pour amortir localement et tout au long du processus (et non en fin de processus) les déformations engendrées par la généralisation d'objets de premier ordre assurant ainsi une bien meilleure qualité dans le maintien des relations entre objets. On notera que ce processus n'est pas limité à une généralisation à but cartographique.



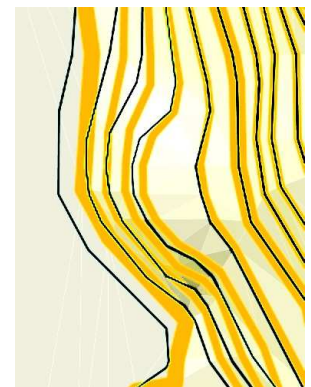
1. Données initiales : des bâtiments au pied d'une pente abrupte.



2. Données finales **sans** le modèle GAEL : les bâtiments se sont déplacés sur la pente.



3. Données finales **avec** le modèle GAEL : les bâtiments ont poussé le relief en se déplaçant.



4. Détail de la déformation appliquée au relief (les courbes initiales sont en noir)

Pour en savoir plus...

Gaffuri J. 2008. « Généralisation automatique pour la prise en compte de thèmes champs: le modèle GAEL. » Thèse de doctorat, université Paris-Est, France.

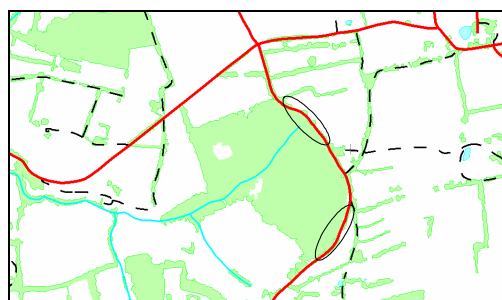
Gaffuri J. 2007. "Field deformation in an agent-based generalisation model: the GAEL model", In *Proceedings of GI-days 2007 - young researches forum*, IFGI prints vol.30, p.1-24.

Contexte et objectifs

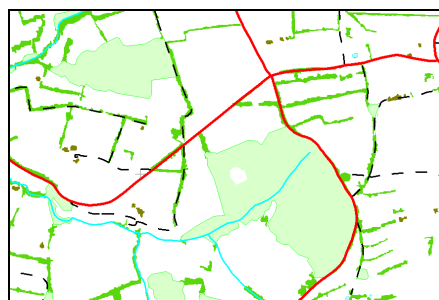
Etude technique réalisée en 2006 à la demande du Service de Bases de données Vecteur de l'IGN. L'IGN et l'Inventaire Forestier National (IFN) mettent en place une collaboration pour la saisie d'un « fond vert commun » (données de végétation), dont sont ensuite dérivées, à l'IGN la couche végétation du Référentiel à Grande Echelle (RGE), et à l'IFN les données forestières après saisie thématique complémentaire. L'étude avait pour objectif 1/ de proposer des outils de traitement du fond vert « brut » (issu de l'extraction semi-automatique sur orthophotos), de manière à faciliter la saisie thématique réalisée à l'IFN, 2/ d'améliorer la cohérence de la végétation avec les autres couches du RGE, et 3/ de faciliter en aval la production de la Carte de Base IGN.

Résultats

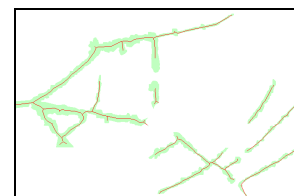
L'étude a permis de proposer des outils 1/ de détection des haies et bosquets, 2/ de recalage des forêts sur les réseaux du RGE et de simplification de leurs contours, et enfin 3/ de squelettisation des haies (reprise des résultats d'un stage de 2005). Ces outils ont été industrialisés par le Projet Nouvelle Carte de Base et sont utilisés en production à l'IGN.



Etat initial : classe « zones arborées » et réseaux routier et hydrographique du RGE. Les ellipses signalent des endroits où la cohérence topologique végétation/réseaux pose problème.



Après traitement (sauf squelettisation des haies). Forêts en vert clair, haies en vert foncé, bosquets en marron. Les forêts sont recalées sur le réseau et simplifiées.



Squelettisation automatique des haies

Pour en savoir plus ...

COGIT 2006. « Généralisation et intégration du fond vert ». Rapport technique, SR/2006-0369, juillet 2006.

PROJET EUROSDR DE TEST DE LOGICIELS DE GENERALISATION DU COMMERCE

Cécile Duchêne, Guillaume Touya

Contexte et objectifs

Projet commencé fin 2006, coordonné par Jantien Stöter sous l'égide d'EuroSDR, impliquant six instituts producteurs de données géographiques (ICC Catalogne, KMS Danemark, IGN Espagne, IGN France, TDK Pays-Bas, OSGB Royaume-Uni) et trois instituts universitaires (Enschede, Hanovre, Zürich). L'objectif est d'évaluer les capacités de quatre logiciels de généralisation commercialisés par quatre industriels volontaires pour ce test : ISpatial, ESRI, l'université d'Hanovre et Axes Systems.

Résultats

Quatre jeux de données test ont été fournis par l'IGN-F, l'OSGB, l'ICC et TDK, accompagnés de spécifications présentées sous forme de contraintes à satisfaire. Les tests ont été réalisés de juin 2007 à avril 2008, d'une part par des testeurs de l'équipe projet, d'autre part par les industriels. Une triple évaluation des résultats est en cours : évaluation visuelle comparative, évaluation de chaque résultat par un cartographe de l'institut ayant fourni le jeu test, évaluation automatique sur quelques critères quantitatifs. Sans anticiper sur les résultats finaux, ce test nous a permis de progresser au niveau méthodologique tout en capitalisant des travaux menés antérieurement au COGIT, notamment sur la définition de spécifications de généralisation et sur l'évaluation de données généralisées.

Pour en savoir plus ...

Stöter J., Anders K.-H., Baella B., Burghardt D., Davila F., Duchêne C., Pla M., Regnauld N., Rosenstand P., Schmid S., Touya G., Uitermark H., 2008. "A study on the state-of-the-art in automated map generalisation implemented in commercial out-of-the-box software." *11th ICA Workshop on progress in automated map generalisation*, Montpellier, 2008, http://aci.ign.fr/montpellier2008/papers/13_Stoter_et_al.pdf.

THEME 3 : SEMIOLOGIE GRAPHIQUE

Rédigé par Elodie Buard

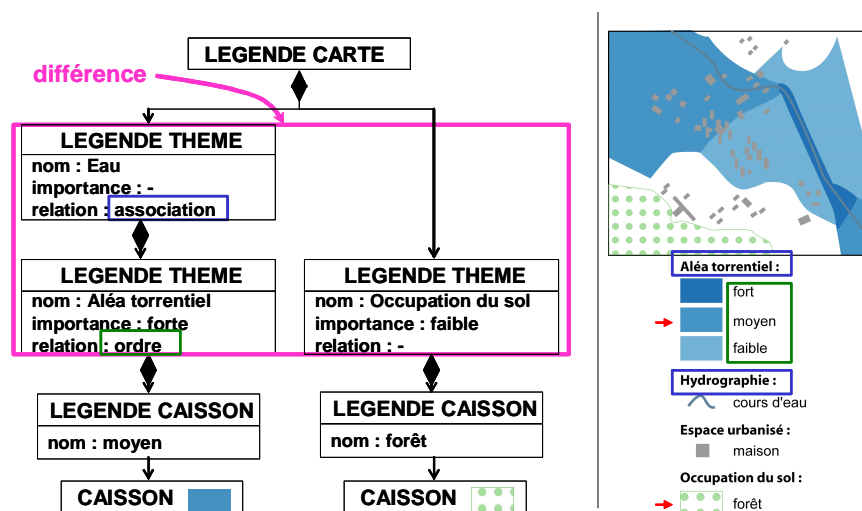
OBJECTIFS ET CONTEXTE

Cette action de recherche vise à améliorer les légendes des cartes. Elle s'inscrit dans la fabrication de 'cartes à la carte' en lien avec l'action de recherche sur l'accès aux services.

Si de nombreuses études s'attachent à résoudre les problèmes techniques de fabrication de cartes, peu d'équipes se penchent sur la conception de légende. Or la légende fait partie des éléments clés pour la compréhension du message contenu dans la carte. Dans un contexte de conception de produits à la demande, l'utilisateur va être amené à choisir non seulement le contenu d'information mais aussi sa symbolisation. La conception d'une légende lisible et facilement interprétable est le métier du cartographe-sémiologue, il s'agit donc ici de formaliser et de mettre à disposition une partie des connaissances cartographiques pour l'amélioration des légendes conçues par ou pour des utilisateurs non cartographes. Ainsi cette action de recherche se propose de concevoir des outils permettant soit d'analyser et d'améliorer automatiquement une légende existante, soit de mettre à disposition des éléments de légende adaptés à des conditions de perception de l'utilisateur, notamment les déficiences visuelles.

Cette action de recherche est présentée dans le Schéma Directeur de la Recherche et des Développements 2006-2010 à l'IGN, qui précise que l'objectif premier est le développement des méthodes permettant d'améliorer la qualité des légendes des produits cartographiques et leur adaptation en fonction de contraintes spécifiques liées à l'utilisateur. Cette action de recherche valorise les produits IGN auprès d'utilisateurs néophytes en leur apportant un service, que ce soit l'amélioration des contrastes de couleurs ou la réalisation de légendes à la carte.

Le travail de thèse d'Elisabeth Chesneau a été à l'origine de cette action de recherche. Sa thèse, soutenue en 2006, traite de l'amélioration automatique des contrastes de couleurs. Le modèle conceptuel de données permet de représenter et de gérer le graphisme et la légende dans un SIG. Chaque ligne de légende possède des relations d'ordre, de différence ou d'association avec les autres lignes de la légende.

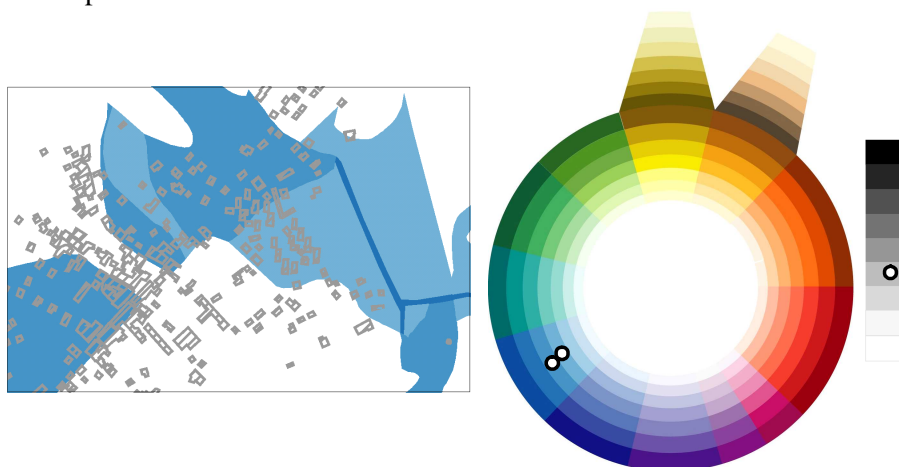


Représentation des lignes de légende « Aléa torrentiel moyen » et « Occupation du sol » dans le modèle de [Chesneau 2006]

Nos recherches utilisent également une palette de couleurs adaptée à la cartographie du risque (à partir de la palette de Cynthia Brewer). Un cercle chromatique a donc été créé, en tant qu'espace de représentation des couleurs des lignes de légendes. Dans ce cercle, une métrique a été proposée de façon à calculer les contrastes entre couleurs en manipulant la clarté et la teinte d'une couleur.

Par ailleurs des règles ont été identifiées pour privilégier des couleurs qui préservent les relations d'ordre, de différence et d'association qui existent entre les thèmes de la légende.

Le processus d'amélioration de légende se traduit opérationnellement par le déplacement des couleurs sur le cercle soit par analyse des contrastes dans la carte (approche d'Elisabeth Chesneau) soit par analyse directe des proximités dans le cercle.



Représentation des couleurs sur un cercle chromatique pour le calcul des contrastes

Le travail de thèse d'Elisabeth Chesneau achevé en 2006, le premier objectif a été de procéder à une réécriture du module sous GeOxygène en améliorant certains éléments du modèle notamment le calcul des contrastes entre couleurs et les méthodes de convergence.

Au niveau international, nous nous sommes principalement inspirés des résultats de recherches produits par Cynthia Brewer, qui a proposé en 2001 un éditeur de couleur (ColorBrewer) ainsi que des légendes types pour des cartes thématiques. Cet éditeur est un outil en ligne d'aide à l'application de couleurs. Il ne s'applique pas à la carte d'un utilisateur mais donne des conseils sur l'utilisation de couleurs, en s'appuyant sur une carte des Etats-Unis, découpée en zones. Des indications y apparaissent sur la qualité de la couleur choisie pour la projection sur un écran, pour l'impression ou même pour les daltoniens.

APPROCHES

Dans l'action de recherche, deux approches se présentent pour traiter l'amélioration automatique des contrastes colorés : soit par la légende, soit par la carte en elle-même.

L'approche par analyse des contrastes dans la légende a été traitée en posant des contraintes aux lignes de légendes, contraintes identifiées par la relation entre les thèmes. Deux approches ont été analysées. La première à base de contraintes (2007) et la deuxième à base d'agents (2008). Ces travaux faits sous forme de stages ont été encadrés par Elodie Buard.

L'approche par analyse des contrastes dans la carte est celle suivie par Elisabeth Chesneau. Les objets analysés sont ainsi les objets cartographiques symbolisés dans la carte et non plus les lignes de la légende. Ainsi chaque objet cartographique calcule ses contrastes avec ses voisins et une agrégation des contrastes est faite au niveau de chaque famille cartographique qui correspond à une ligne de légende.

Cette deuxième approche est plus précise que la première puisque seuls les contrastes réels sont calculés. Par contre elle nécessite beaucoup plus de calculs puisqu'elle intègre toutes les relations de voisinage entre les objets de la carte.

Ces deux approches nécessitent des connaissances sur les couleurs. Une palette de couleurs définie par Elisabeth Chesneau est utilisée en permanence dans cette action de recherche. Cette palette combine des couleurs pures et des couleurs grisées, pour 134 couleurs au total.



Palette de couleurs utilisée dans l'action de recherche

En 2006 le calcul des valeurs de contraste n'était pas très satisfaisant. Entre 2006 et 2008 Elodie Buard s'est penchée sur l'amélioration de la métrique de couleurs par acquisition de connaissances en effectuant des tests auprès d'experts.

PRINCIPAUX RESULTATS

Amélioration par analyse de la légende

Deux méthodes ont été implémentées : une à base de résolution de contraintes en 2007, l'autre à base d'agents. La première méthode permet effectivement de corriger des couleurs mais à tendance à produire le même résultat quelle que soit la légende initiale.

La deuxième approche à base d'agents est en cours de finalisation. Les agents couleurs se déplacent les uns des autres pour respecter au mieux les relations d'ordre, d'association et de différence. Ainsi la légende initiale est préservée au mieux.

Amélioration par analyse des contrastes dans la carte

Le modèle d'Elisabeth Chesneau d'amélioration de la légende par analyse des contrastes dans la carte a été réimplémenté sous GeOxygène. Les cartes sont analysées à la volée, de façon automatique. Si le système s'appuie sur une palette de couleurs, la légende peut contenir des couleurs absentes de cette palette. Le système détecte automatiquement les relations entre les thèmes d'après l'organisation de la légende et prend la couleur de la palette la plus proche de celle de l'utilisateur. Les cartes initiales et les cartes après traitement sont visibles sous Open Jump. Les couleurs sont changées étape par étape en appuyant sur un bouton qui réitère le processus de changement. L'arrêt automatique et la convergence en général sont en cours d'étude.

Les contrastes entre couleurs

En parallèle, la métrique des couleurs a été améliorée par des tests auprès d'experts cartographes ; Laurence Jolivet a développé en 2006 des outils d'acquisition de connaissances en utilisant un module utilisé dans l'action de recherche en généralisation (module MAACOL). Ce module a permis de simplifier le processus d'acquisition des connaissances et d'avancer nos réflexions sur la conception d'un test pertinent sur les couleurs.

En 2007 un protocole de test a été écrit par Elodie Buard pour améliorer le calcul des contrastes de teinte et de valeur dans le cercle chromatique. Le test a été appliqué aux couleurs pures du cercle (cercle principal). Les tests ont permis de calculer de meilleures valeurs de contraste de teinte et de clarté. Ils ont aussi mis en évidence le déséquilibre du cercle chromatique. Pour illustrer ces résultats, nous avons déformé ce cercle (voir pages suivantes).

En 2008 de nouveaux tests ont été réalisés pour mieux situer les couleurs ocre et marron, entre elles et par rapport au cercle des couleurs pures. Ainsi nous avons pu compléter notre matrice de calcul de contrastes et comparer nos résultats avec ceux obtenus en utilisant le système TSL (Teinte Saturation Luminosité). Les tests montrent que les résultats sont dans l'ensemble similaires (cohérence) sauf pour certains couples de couleurs où le contraste perçu (celui que nous avons calculé) est bien moins fort que celui calculé par le système TSL. C'est le cas par exemple du couple de couleurs bleu-violet.

L'amélioration des mesures de contraste entre couleurs permet d'améliorer les notes de contraste calculées dans le module d'amélioration automatique de la légende et donc d'assurer une meilleure convergence.

Contexte et objectifs

Travaux de deux mois réalisés par des stagiaires en 2006 et 2007, puis finalisés par Elodie Buard en 2008. Ces travaux consistent à identifier le meilleur système de représentation des couleurs pour une amélioration cartographique et à définir une métrique des couleurs pour calculer les contrastes. Les mesures de contrastes proposées par Elisabeth Chesneau dans sa thèse ont été améliorées en réalisant des tests auprès de cartographes.

Résultats

Laurence Jolivet a tout d'abord mis en place un test de contrastes colorés sous Lamps2, grâce à un module permettant 1/ de présenter à un cartographe expert des couples de couleurs et 2/ d'enregistrer des valeurs de contrastes qualitatives, de teinte ou d'ordre. Quelques couples de couleurs du cercle chromatique ont été évalués, sur des critères de teinte et de clarté.

Elodie Buard a réévalué la métrique des couleurs par des tests plus poussés sur le cercle de couleurs pures. Ces tests ont montré l'hétérogénéité du cercle chromatique, tant pour la teinte que pour la clarté. Le cercle a ainsi été déformé pour représenter les distances perçues, en teinte et clarté, entre les couleurs. Les écarts entre couleurs ont été calculés de façon statistique à partir des résultats des tests.

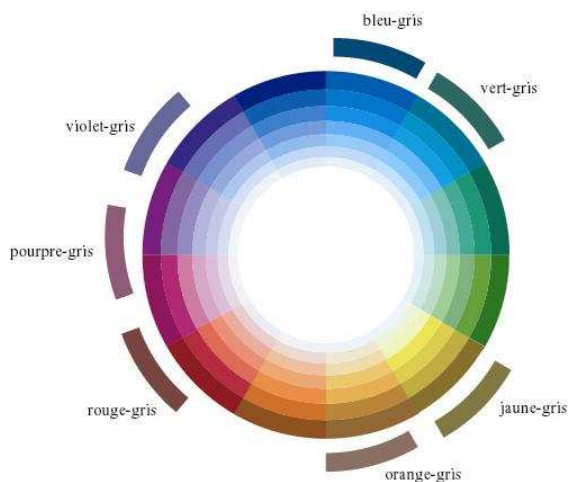


Cercles de teinte et de clarté déformés

Nicolas Gonindard a évalué les contrastes colorés pour les gris, le marron et l'ocre.

Le premier résultat est d'avoir comparé des valeurs théoriques de teinte et de clarté (valeurs radiométriques) aux valeurs perçues, testées auprès d'experts cartographes. Cette analyse étudie l'équilibre du cercle et valide les déformations précédentes du cercle chromatique.

Ensuite, les cercles ont été reliés entre eux, par une projection sur le cercle de couleurs pures. Les cercles de gris semblent être eux aussi déséquilibrés, puisqu'il n'y a pas d'écart constant entre les teintes grisées. L'écart entre les teintes et les projections entre cercles a permis de déterminer de meilleures valeurs de contrastes de teinte et de clarté.



Projection des couleurs grisées sur le cercle de couleurs pures

Pour en savoir plus ...

Jolivet L., 2006, « Analyse des contrastes de couleurs - Mise en place d'un test sous le logiciel Lamps 2 », rapport de stage, soutenu le 6 mars 2006, Master 2 Carthagéo-pro, ENSG, Paris I, Paris VII

Buard E., Ruas A., 2007, "Evaluation of colour contrasts by means of expert knowledge for on-demand mapping", 23rd ICA conference, 4-10 August 07, Moscow.

Gonindard N., 2008, « Tests d'évaluation des contrastes colorés », rapport de stage, soutenu le 7 mars 2008, Master 2 Carthagéo-pro, ENSG, Paris I, Paris VII

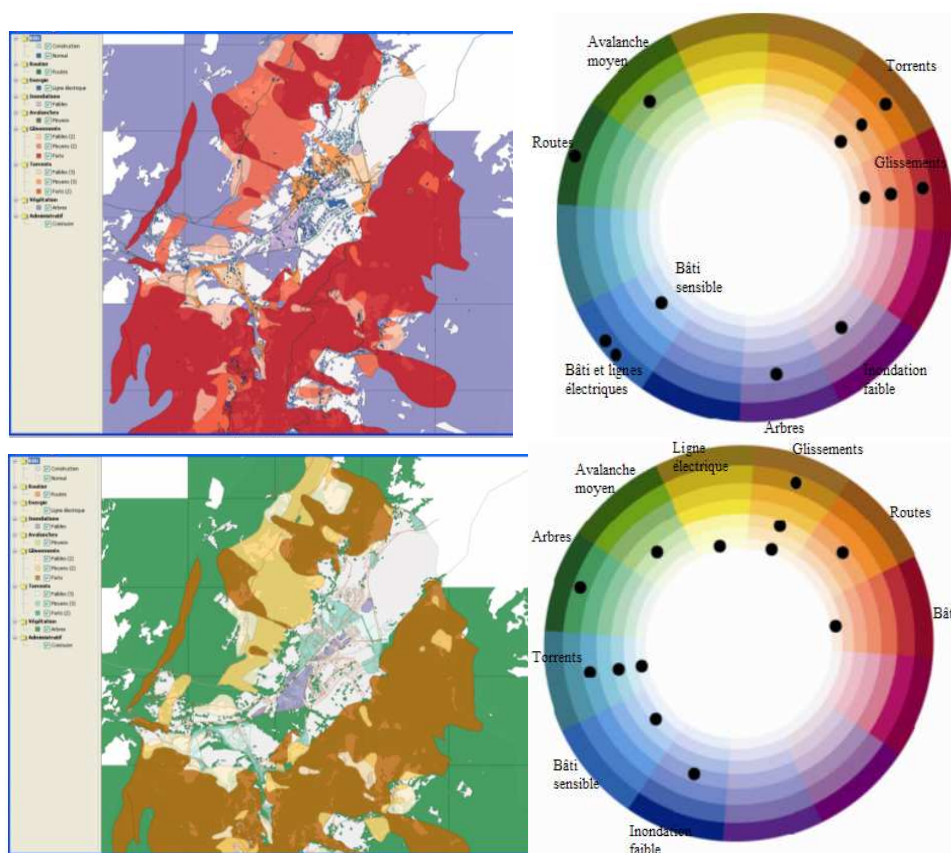
Contexte et objectifs :

Travaux réalisés autour de deux stages de 5 mois en 2007 et 2008, encadrés et finalisés par Elodie Buard. Il s'agit de concevoir une méthode d'amélioration des contrastes, en exploitant uniquement les données de la légende, et en particulier l'adéquation entre les couleurs initiales des thèmes de la légende et les relations entre les thèmes, à savoir les relations d'association, de différence et d'ordre.

Résultats

Benjamin Ganem a développé un système de résolution de contraintes, les contraintes étant la position des couleurs sur le cercle et leur placement par rapport aux autres couleurs.

La figure suivante montre une carte avant (en haut) et après (en bas), et les cercles à côté correspondent aux couleurs des lignes de légendes dans chacun des états. La légende après traitement est plus lisible que la première : les thématiques comme les arbres en vert ont été respectées et il n'y a plus de problème d'incohérence entre le choix des couleurs et les relations entre thèmes.



Carte initiale en haut, carte finale en bas

Cependant, ce processus donne pratiquement toujours les mêmes légendes de sortie, parce que les contraintes sont trop fortes et que la légende initiale n'est pas suffisamment prise en compte.

Ainsi *Salima Taibi* en 2008 a implémenté un système multi-agent où chaque agent-couleur bouge de façon autonome et en communiquant avec les autres agents. En fonction des relations à respecter, l'agent va se déplacer sur le cercle soit pour se rapprocher d'un autre agent, soit pour s'en éloigner. Les déplacements peuvent se faire selon la teinte (déplacement circulaire) ou la clarté (déplacement radial). Ce type de méthode permet une bien meilleure prise en compte de la légende initiale.

Les travaux actuels visent à finaliser cette approche.

Pour en savoir plus ...

Taibi S., 2008, « Système multi-Agents sur un Cercle Chromatique » Master 2 informatique, Université Paris 5 (René Descartes) spécialité Systèmes Multi-Agents Stage soutenu le 29 juin 2008

Ganem B., 2007, « Système multi-agents sur un cercle chromatique », rapport de stage, soutenu le 20 juin 2007, Master 2 informatique, Université Paris 5 (René Descartes) spécialité Systèmes Multi-Agents, 88 pages

Contexte et objectifs

Ce travail concerne l'implémentation sous GeOxygene du module d'analyse et d'amélioration des contrastes colorés d'une carte à partir des objets cartographiques. Il s'agit donc de réimplémenter et de compléter les résultats de la thèse d'Elisabeth Chesneau.

L'objectif du module est d'automatiser le processus d'amélioration des légendes utilisateur en conservant l'originalité de la légende initiale proposée par l'utilisateur.

Cette implémentation a eu lieu de 2007 à 2008 et a nécessité au préalable le calcul des matrices de teinte et de clarté par les tests sur les contrastes colorés présentés précédemment, et leur intégration dans le module.

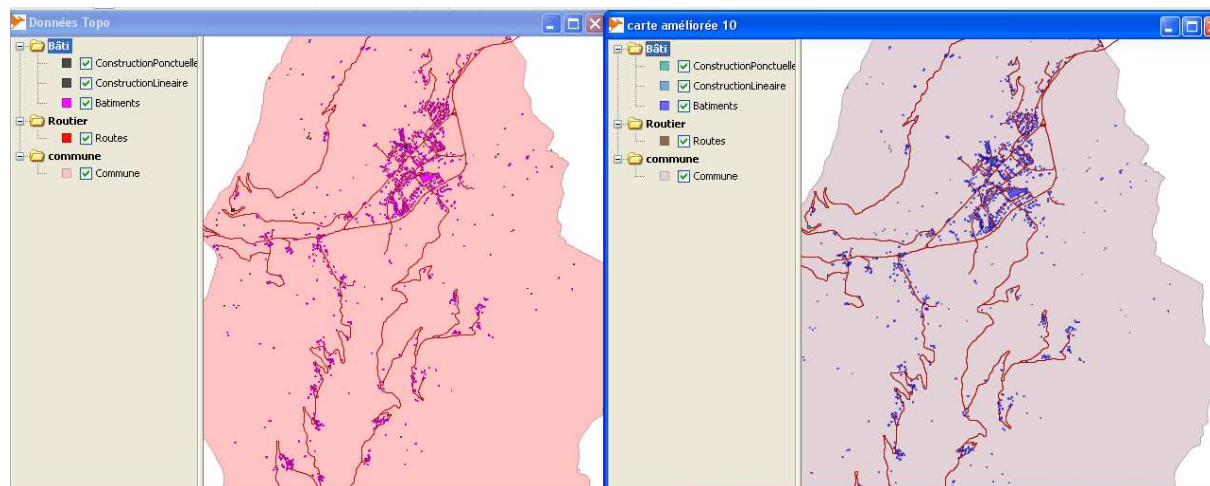
Résultats

Le module implémenté analyse une carte avec une légende initiale, en changeant les couleurs de façon progressive.

L'étape *d'analyse des contrastes de couleurs* prend en compte 1/ les relations entre thèmes de la légende et 2/ les contrastes entre les objets cartographiques voisins sur la carte (voit-on suffisamment cet objet ?) Les contrastes sont calculés à partir de matrices de contrastes de teinte et de clarté (voir pages précédentes).

L'étape de *changement de couleurs* fait intervenir également l'impact du changement de couleurs du thème en fonction de son nombre d'objets et de sa surface sur la carte.

Le processus est itératif. Les couleurs sont modifiées à chaque itération en se déplaçant sur le cercle chromatique vers des couleurs voisines, jusqu'à satisfaction de la légende ou non progression.



Carte initiale à gauche, carte au bout de 10 itérations du module d'amélioration à droite

Les travaux actuels concernent la convergence du système. Il s'agira ensuite de comparer la qualité des légendes selon les différentes méthodes disponibles au laboratoire. Ce module sera déposé un OpenSource sur GeOxygène courant 2009.

Pour en savoir plus ...

Chesneau E., 2006, « Modèle d'amélioration automatique des contrastes de couleurs en cartographie - Application aux cartes de risque, » Thèse de doctorat, Université de Marne La Vallée, soutenue le 23 mai 2006

THEME 4 : ACCES AUX SERVICES ET AUX DONNEES

Rédigé par Catherine Dominguès et Bénédicte Bucher, responsable de l'action de recherche

CONTEXTE ET OBJECTIFS

Le recours à la carte pour représenter une information géo-référencée est de plus en plus fréquent et facile. Dans le même temps, l'accès aux données géographiques, leur représentation et leur manipulation évoluent avec les transformations survenues dans le domaine de la diffusion d'information.

En terme d'évolutions récentes qui impactent notre action de recherche, nous pouvons citer :

- La directive européenne INSPIRE (<http://inspire.jrc.it>) vise à organiser et faciliter la diffusion en Europe d'une information géographique utile à la définition, au suivi et à l'évaluation des politiques européennes, en particulier dans le domaine de l'environnement. La mise en place d'une politique commune repose sur l'interopérabilité et la réutilisabilité des bases de données produites par les états.
- Le GéoPortail (<http://www.GeoPortail.fr>) est le portail géographique français de service public. Il permet de visualiser gratuitement des données géographiques de référence et des informations localisables fournies par des services publics.
- Les logiciels de création de cartes, gratuits ou presque, souvent OpenSource, sont de moins en moins complexes à utiliser.
- Des sites web proposent des services de cartographie gratuits. De nouveaux serveurs de cartes conformes aux spécifications Web Mapping Service (WMS) de l'Open Geospatial Consortium (OGC) [OGC 05] apparaissent. Des clients qui permettent d'interroger simplement les sites respectant les spécifications OGC et d'importer des données géographiques distantes sont également proposés gratuitement.

Cependant, l'adaptation des données au contexte de l'utilisateur reste un problème difficile. Afin que l'utilisateur dispose, pour le territoire qui l'intéresse, d'informations géographiques dont il puisse évaluer la validité et la pertinence, des services doivent être développés : *extraction de données, intégration de données provenant de sources diverses, visualisation cartographique, analyse spatio-temporelle.*

Une fois les problèmes techniques résolus, le besoin d'expertise demeure : tout détenteur d'un ordinateur et d'un accès internet a la possibilité de faire une carte pour servir son objectif, mais sans compétences spécifiques, il risque de produire une carte fautive (parce qu'il a mal compris la signification des données qu'il a choisies), inefficace par rapport à son objectif (parce qu'il a mal représenté les données qu'il a choisies), ou stéréotypée et non adaptée à ses préférences esthétiques.

Dans ce contexte, cette action de recherche a pour objectifs de développer et diffuser des méthodes permettant à un utilisateur non spécialiste d'accéder à des données et à des processus de manipulation et d'affichage de ces données. Ces processus (restructuration de données pour dériver des informations géographiques implicites, mise à jour et contrôle qualité des données, rédaction d'une carte, évaluation de cette carte) sont classiquement fournis par l'IGN mais nous souhaitons les adapter aux besoins des utilisateurs. L'action de recherche ne vise pas à concevoir l'ensemble de ces traitements mais à en faciliter l'accès qui est difficile ou incertain pour un utilisateur non spécialiste. L'enchaînement de différents traitements nécessite aussi une expertise que nous souhaitons offrir à un utilisateur non spécialiste.

Nous souhaitons privilégier des accès sur le Web, via un navigateur ou des services Web.

Le développement de méthodes d'accès aux données et à leur manipulation s'appuie sur des outils que nous devons créer : modèles de métadonnées, méthodes de consultation et d'acquisition des métadonnées, interfaces utilisateurs conviviales, etc.

L'action de recherche s'est concentrée sur quelques thématiques peu traitées (et sur les outils correspondants) et dont nous pensons que la mise en place nécessite l'expertise d'un producteur de données ou d'un concepteur de cartes.

Ces thématiques sont les suivantes :

- l'accès à des données et à leur structuration : nous voulons permettre à un utilisateur non expert de changer simplement de représentation pour adapter des données à son contexte (ses concepts métiers et sa plate-forme d'implémentation) ;
- les outils de manipulation des données et les connaissances sur ces manipulations ;
- les interactions homme-machine dans le but de construire une légende ;
- la conception d'outils génériques permettant de faciliter le déploiement de méthodes existantes en tant que services web ;
- la définition et l'organisation des connaissances cartographiques nécessaires pour décrire et concevoir une carte adaptée au besoin d'un utilisateur.

APPROCHE

La démarche que nous avons adoptée garantit que les travaux s'intègrent dans une même plate-forme de consultation, au plan conceptuel voire opérationnel.

L'intégration conceptuelle de l'ensemble des travaux de recherche est fondée aujourd'hui sur l'utilisation de deux modèles :

1. le modèle intégré de métadonnées,
2. le modèle conceptuel de carte.

La plate-forme de programmation GeOxygène fournit un cadre de stockage et de manipulation des données géographiques. Elle contribue à la cohérence de la démarche de l'action de recherche et garantit la conformité aux standards de l'ISO et OGC.

Cette démarche s'articule autour de quatre thèmes de recherche :

1-Accès aux données et à leur structuration

L'utilisateur peut souhaiter adapter son jeu de données à son application, c'est-à-dire le simplifier, choisir les noms de ses classes, filtrer l'information inutile, ou dériver celle qui y est contenue implicitement. Pour faciliter et sécuriser cette adaptation, les opérations s'appliquant aux données doivent être déduites automatiquement des transformations du schéma, spécifiées par l'utilisateur. Cette déduction nécessite le paramétrage et l'enchaînement automatiques d'opérations élémentaires qui peuvent être géographiques (par exemple, la création des objets d'une nouvelle classe par agrégation d'objets d'une classe initiale) ou non. Dans le cadre de sa thèse (soutenue en septembre 2007) Sandrine Balley a étudié les métadonnées décrivant les données (en particulier les métadonnées décrivant leur structure) et les opérations de transformations de structure qui s'appuient sur ces métadonnées, et proposé une interface graphique permettant à l'utilisateur de paramétrer sa transformation sur le schéma externe des données.

2-Accès aux manipulations de données

Bénédicte Bucher a étudié la mise en ligne de processus de dérivation qui s'appuient sur GeOxygène et sur des services Web. Elle a construit un modèle de métadonnées ainsi qu'une application de catalogage fondée sur ce modèle. L'adaptation au contexte est limitée à l'adaptation aux données de l'utilisateur. Elle est traitée sous la forme du paramétrage de processus spécifiques (prétraitements). Laurence Jolivet s'est concentrée sur un processus de dérivation particulier : la conception d'une carte sur mesure. Ce processus s'appuie sur le module d'amélioration automatique de légende, étudié par l'action de recherche LUCIL, afin d'adapter la carte produite aux besoins et aux goûts de l'utilisateur.

3-Interactions homme-machine

Sidonie Christophe propose une solution pour aider l'utilisateur à exprimer sa créativité tout en l'aidant à concevoir une carte efficace. Son approche est fondée sur la construction d'un dialogue collaboratif entre un serveur cartographique et un utilisateur. Ses travaux reprennent le principe introduit dans le travail de thèse de Frédéric Hubert (thèse soutenue au COGIT en juillet 2003) : au lieu de fonder la communication exclusivement sur un échange en langage naturel, le système expérimente d'autres types de dialogue. Sidonie Christophe propose à l'utilisateur de réagir à des exemples de cartes ou de palettes de couleurs qui pourraient s'appliquer à son jeu de données. Grâce à cette approche, il n'est pas

nécessaire que l'utilisateur possède ou acquière des connaissances en cartographie ou en sémiologie graphique pour pouvoir définir les caractéristiques de la carte qu'il veut élaborer.

4-Construction de ressources

Pour concevoir une carte sur mesure, il faut 1/ comprendre le besoin de l'utilisateur et disposer des connaissances nécessaires pour l'interpréter, puis 2/ construire une carte efficace et adaptée à ce besoin. Constatant que leurs différents travaux s'appuyaient sur des connaissances très parcellaires et non formalisées, les chercheurs ont décidé de construire des connaissances partagées qui pourraient servir de support aux travaux d'adaptation au contexte de l'utilisateur.

Deux types de corpus ont été identifiés et construits :

1. un corpus d'échantillons cartographiques. Ces échantillons constituent de bons exemples sur lesquels l'utilisateur peut fonder des comparaisons, des sélections, des exclusions. Il peut ainsi exprimer ses goûts et ses préférences sans avoir besoin d'acquérir les connaissances spécifiques en cartographie ou en sémiologie graphique ;
2. un corpus textuel. Catherine Domingues exploite un corpus d'environ 150 000 mots. Ces textes proviennent de sources variées (manuels, cours en ligne, articles de journaux) et rendent compte de différents points de vue sur la carte, en particulier ceux de l'utilisateur néophyte et du concepteur expert. L'objectif est de déterminer le vocabulaire et les concepts pertinents pour spécifier un besoin de carte. Ce corpus a été mis en forme afin d'être exploité avec des techniques de traitement automatique du langage naturel.

PRINCIPAUX RESULTATS

Les travaux de l'action de recherche ont conduit à définir des modèles et à programmer des applications qui s'appuient sur des ressources partagées.

Modèles

Le travail de thèse de Sandrine Balley l'a menée à définir un premier modèle permettant de représenter la structure d'un jeu de données de façon cohérente à tous les niveaux d'abstraction. Il a été mis au point en enrichissant et reliant des modèles standards de l'OGC.

Le second modèle lui a permis de formaliser le processus de restructuration pour faciliter son automatiser.

Dans sa thèse Yann Abd-el-Kader a proposé un modèle de métadonnées utile au catalogage de traitements géographiques. Il a associé à ces métadonnées une base de règles permettant d'adapter la description d'un traitement à la plate-forme et aux compétences de l'utilisateur. Bénédicte Bucher a combiné ce modèle avec celui sur les manipulations d'applications complexes issu de ses propres travaux de thèse.

Le modèle de dialogue proposé par Sidonie Christophe permet de décrire les méthodes associées à l'aide à la conception de légende : stratégies pour faire s'exprimer l'utilisateur sur ses préférences ; méthodes de représentation des connaissances, de traduction en contraintes, de résolution de ces contraintes, d'évaluation et de retouche.

Un modèle de description formelle de carte a été proposé par Catherine Domingues dans le but d'ordonner une base d'échantillons cartographiques. La description des échantillons permettra d'interroger la base sur des critères pertinents ; en particulier, la proposition d'échantillons cartographiques (dans l'application de Sidonie Christophe) adaptés aux contraintes et aux goûts de l'utilisateur pourra utiliser cette description.

Applications

Laurence Jolivet a développé des modules d'aide à la création de cartes sur mesure sur Internet. Il s'agit d'un ensemble de services Web qui acceptent en entrée des spécifications décrites par un

utilisateur, et qui font appel à des fonctionnalités d'expertise cartographique. L'objectif consiste à proposer un contenu et une symbolisation adaptés aux besoins des utilisateurs. Trois services ont été identifiés afin d'accompagner les étapes de la création de cartes.

1. Le premier service a pour but de donner une description formelle de la carte souhaitée à partir de la zone géographique, du type de carte, etc.
2. Le deuxième définit une légende pertinente en fonction de cette description formelle et des choix de l'utilisateur.
3. Enfin, une évaluation des contrastes colorés de la légende est calculée via un dernier service afin d'améliorer la lisibilité de la carte.

Cette application intègre des connaissances et des modules développés dans les actions de recherche SISSI et LUCIL : sélection automatique de données à partir de métadonnées, corpus de connaissances cartographiques, analyse d'échantillons de cartes, travaux sur la perception des couleurs.

Une application de dialogue permettant à un utilisateur d'être assisté dans sa démarche de conception de légende a été développée par Sidonie Christophe. Le but de ce dialogue est de concevoir une légende correcte (par rapport aux règles de sémiologie graphique) et satisfaisante (par rapport aux goûts de l'utilisateur). Le système apporte l'expertise cartographique qui peut faire défaut à l'utilisateur, lui propose des solutions variées, évalue la qualité de la légende et le laisse libre de choisir la légende qui lui convient le mieux et de la retoucher si nécessaire. L'utilisateur évalue et commente les propositions du système, il peut refuser, recommencer, changer de méthode d'approche. L'application de dialogue permet donc à l'utilisateur d'être créatif tout en concevant une carte cartographiquement correcte.

Ressources

La construction de ressources partagées a débuté par la création de deux corpus :

1. un corpus d'échantillons cartographiques (cf. stages de Pierre Philippe et Jérémy Renard). Ce corpus constitue un support pour tous les travaux qui utilisent des exemples de cartes. Un modèle de description formelle de carte a été ébauché. La description à l'aide de ce modèle d'un nombre d'échantillons significatifs sera prise en charge par l'ensemble des membres de l'action de recherche.
2. un corpus textuel (cf. travail de Catherine Domingues). A partir de ce corpus, il s'agit d'organiser et hiérarchiser les termes appartenant aux domaines spécifiques de la cartographie et de la sémiologie graphique dans une ontologie.

Cette ontologie garantira que les concepts et les termes qui les décrivent ont des acceptions communes pour l'ensemble des membres de l'action de recherche. Elle servira de support à la définition de règles d'inférence (concernant la description du besoin, la construction de la légende, les règles de base de sémiologie graphique) pour la spécification du besoin de l'utilisateur et la création d'une carte adaptée. Tous les concepts manipulés par ces règles (à écrire) devront figurer dans l'ontologie. Elle constituera ainsi un réservoir de concepts pour la rédaction des règles d'inférence.

Perspectives

Les résultats obtenus ont montré plusieurs pistes d'amélioration :

- Nous souhaitons mieux modéliser le contexte de l'utilisateur et pour cela enrichir les règles d'inférence qui permettent de déduire, à partir de la description produite par l'utilisateur, des informations pertinentes pour améliorer les propositions de légendes.
- Les interactions entre l'utilisateur et le système sont fondées sur des propositions d'échantillons, d'exemples, de palettes de peintres. Nous pensons améliorer les stratégies de proposition de cartes en articulant et en combinant ces différentes propositions.
- Nous souhaitons élaborer un prototype d'une application de "Carte à la demande" qui privilégie la réutilisabilité du code et des ressources. Elle s'appuiera sur la plate-forme logicielle GeOxygène ; elle combinera les ressources et les modules logiciels développés par l'action de recherche (catalogage des traitements et exploitation des métadonnées, extraction de données, spécification du besoin de carte, spécification de la carte à construire, proposition automatique de symbolisations adaptées) et pourra bénéficier d'autres contributions (par exemple, l'ontologie des objets géographiques GeOnto). L'application de Sidonie Christophe "Cartographie créative à la

demande" permettra de tester des stratégies de conception collaborative de légendes à partir d'exemples. Ces stratégies pourront ensuite être intégrées au prototype de "Carte à la demande".

- Une application de consultation via un navigateur ou des services Web doit être développée. Cela nécessitera de traiter plusieurs problèmes non résolus par les métadonnées comme la gestion des identifiants des objets avant et après traitement.

Contexte et Objectifs

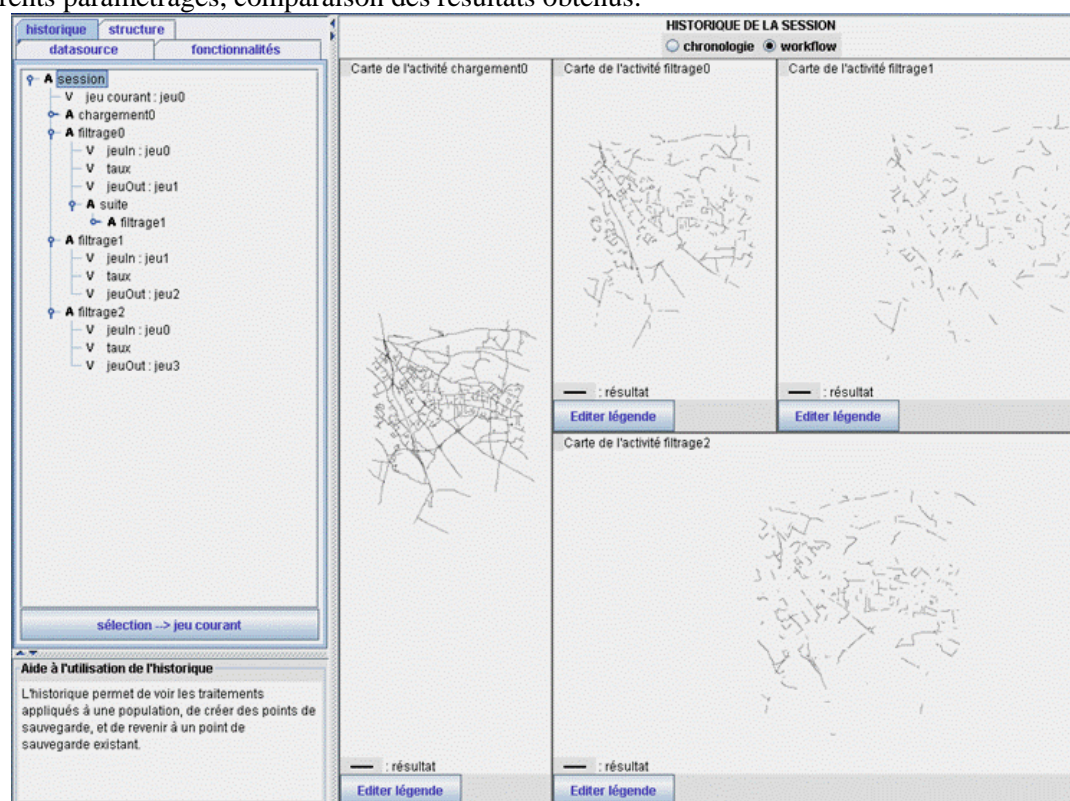
L'utilisation de données géographiques nécessite des outils spécifiques. Avec les progrès de la diffusion de données géographiques, il est de plus en plus nécessaire de donner accès également à l'application de traitements parfois complexes sur ces données. La directive INSPIRE demande ainsi le développement de services d'intégration ou d'analyse spatio-temporelle.

Les objectifs de ces travaux sont de permettre la mise en ligne de méthodes logicielles et du savoir-faire associé pour le paramétrage, la composition et l'interprétation des résultats. Nous nous concentrons sur des méthodes disponibles sur la plate-forme GeOxygène.

Résultats

En juillet 2006, Yann Abd-el-Kader a soutenu sa thèse sur un modèle de métadonnées pour le catalogage de traitements géographiques. Il a associé aux métadonnées une base de règles permettant d'adapter la description d'un traitement à la plate-forme et aux compétences de l'utilisateur. Il a testé l'acquisition de descriptions auprès de développeurs du COGIT.

A la suite de ces expériences, Bénédicte Bucher a intégré ce modèle de métadonnées sur les traitements logiciels avec celui sur les manipulations complexes issus de ses propres travaux de thèse. En tenant compte des retours utilisateurs, elle a développé un module d'édition accessible depuis l'environnement GeOxygène et a conçu un module de consultation sur le Web qui déploie les ressources décrites (méthodes et séquences d'appels à méthodes) sous forme de services invocables depuis un client Web. Ce client a été conçu pour faciliter l'exploration de différentes méthodes et différents paramétrages, comparaison des résultats obtenus.



Client permettant d'explorer les fonctionnalités disponibles sur GeOxygène, de sélectionner, paramétrer et invoquer une méthode et de comparer les résultats obtenus avec différents paramétrages.

Pour en savoir plus...

Abd-el-Kader Y., 2006, « Conception et exploitation d'une base de métadonnées de traitements informatiques, représentation opérationnelle des connaissances d'expert - Application au domaine géographique » Thèse en informatique de l'Université de Caen, Saint Mandé

Bucher B., 2007, "A cartographic Web client to explore remote GI methods", in proceedings of the 23rd ICA conference, Moscow, Russia

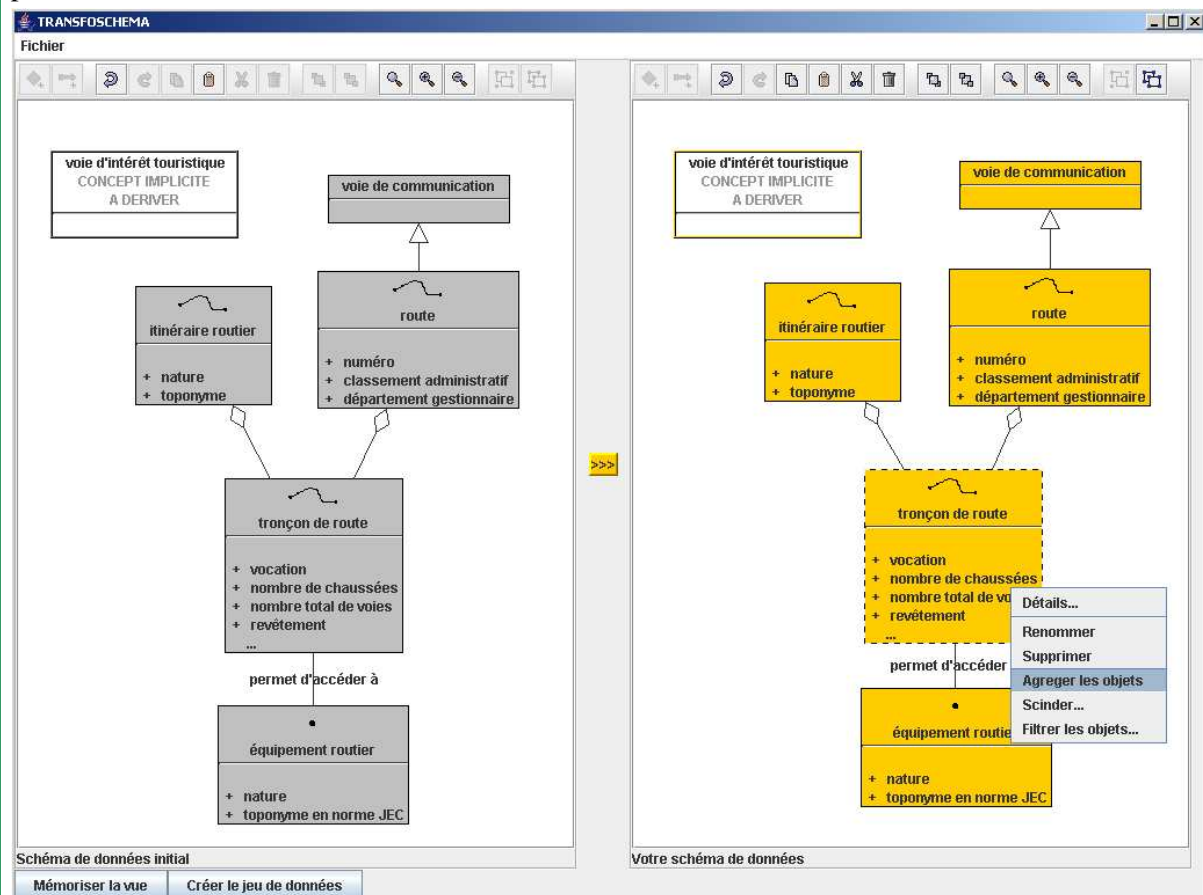
Contexte et Objectifs

Ce travail mené entre 2003 et 2007, dirigé par Thérèse Libourel du LIRMM et encadré au Cogit par Bénédicte Bucher, a reçu le deuxième prix de thèse du concours SIGMA

L'objectif de cette recherche est de développer une méthode et un système permettant à un géomaticien non expert de la gestion de données de personnaliser le contenu et la structure d'un jeu lors de la commande. L'utilisateur paramètre les transformations en manipulant un schéma conceptuel sur une interface graphique et le système les exécute en fin de session. Le système intègre l'expertise du producteur de données, notamment pour souligner l'information implicite dérivable, éviter d'endommager les données et garantir leur mise à jour ultérieure.

Résultats

Cette méthode a été mise en œuvre sur la plate-forme GeOxygène. Elle repose sur deux modèles. Le premier permet de représenter la structure d'un jeu de façon cohérente à tous les niveaux d'abstraction. Il a été mis au point en enrichissant et reliant des modèles standard de l'OGC. Le second formalise le processus de restructuration et facilite son automatiser. Une bibliothèque de 15 transformations de schéma est proposée (scission de classes, suppression d'attributs, etc.) et peut être enrichie par le producteur



Interface de l'application Web. A gauche, le schéma initial des données du producteur. A droite, le schéma en cours de transformation par l'utilisateur. L'accès aux données par leur schéma conceptuel libère l'utilisateur des questions techniques liées à la complexité des données et des plates-formes

Pour en savoir plus...

Balley S., 2007, « Aide à la restructuration de données géographiques sur le Web - Vers la diffusion à la carte d'information géographique ». Thèse en informatique de l'Université de Paris-Est Marne-la-Vallée, Saint Mandé, septembre 2007

Contexte et Objectifs

Thèse de doctorat en sciences de l'information géographique de l'Université Paris Est, sous la direction d'Anne Ruas (IGN-COGIT) et l'encadrement de Bénédicte Bucher (IGN-COGIT).

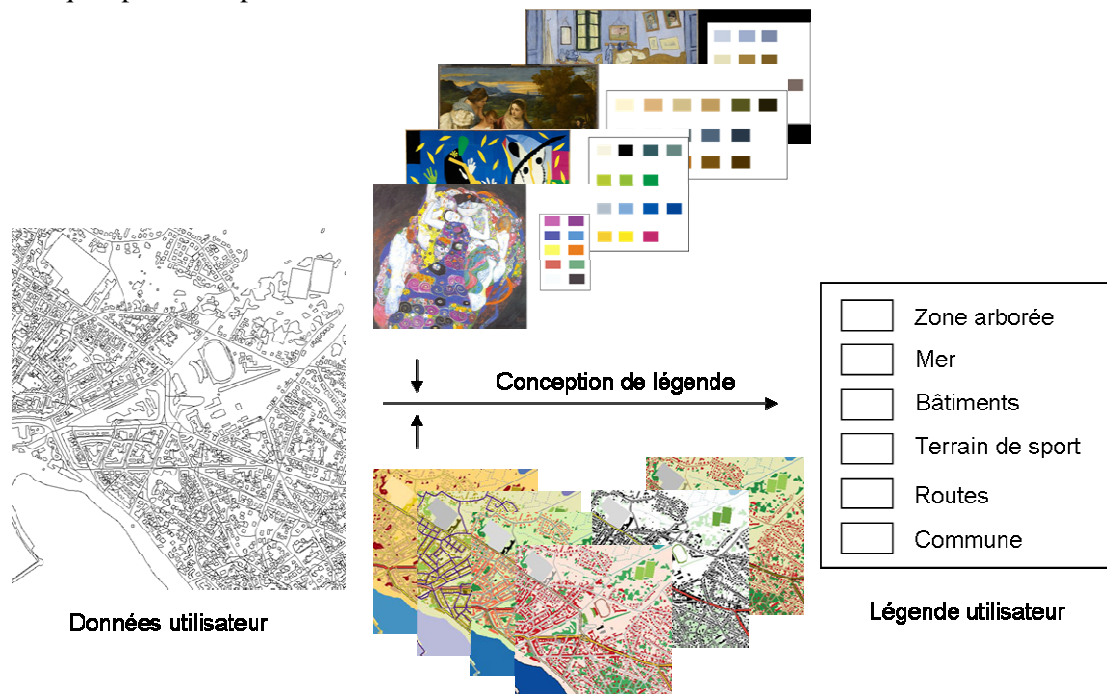
Les outils cartographiques permettant de concevoir des cartes « à la demande » fleurissent sur le Web. Cependant la plupart ne permettent pas de modifier les styles, ni d'évaluer l'efficacité de la future carte. Il est nécessaire d'avoir une bonne connaissance technique de ces outils pour les manipuler aisément et concevoir une belle carte. L'objectif de ce travail est de proposer un modèle d'aide à la conception de légendes qui soient cartographiquement correctes et satisfaisantes pour l'utilisateur.

Approche

Le modèle proposé est un moteur de dialogue entre un utilisateur et le système dont le but est de concevoir une légende correcte et satisfaisante. Ce moteur se décompose en deux parties :

1. Un contrôleur de dialogue gère le bon déroulement du dialogue ainsi que l'état d'avancement de la légende, le respect des contraintes cartographiques et de l'utilisateur, et la satisfaction finale de l'utilisateur. Le contrôleur gère les réponses à l'utilisateur sous forme de stratégies de conception.
2. Des stratégies de conception simulant des démarches possibles utilisées par des cartographes pour concevoir des légendes : deux stratégies sont proposées :
 - *Sélectionner et retoucher un échantillon* : Le système s'appuie sur une base d'échantillons de cartes et cherche à faire choisir la légende la plus satisfaisante à l'utilisateur : elle est ainsi appliquée à ses données. Celui-ci peut ensuite faire des retouches sur cette légende.
 - *Palettes de peintres* : Des palettes extraites de peintures célèbres ont été conçues afin de proposer des couleurs originales et des combinaisons harmonieuses de couleurs. L'application des couleurs à chaque thème se fait selon une base de connaissances cartographiques, et sur des règles d'utilisations conventionnelles des couleurs et des règles de contraste entre couleurs.

Ces processus permettent d'obtenir des légendes correctes et variées. L'utilisateur sélectionne la légende qu'il préfère et peut réaliser des retouches s'il le souhaite.



Conception de légendes à base d'échantillons cartographiques et de palettes de peintres

Pour en savoir plus...

Christophe, S., 2008. « Concevoir des légendes topographiques à l'aide de palettes de peintres », Assemblée Générale du Comité Français de Cartographie (CFC), Jeudi 22 Mai 2008, Musée de la Marine, Paris.

Christophe S., 2007. "Legend design by a novice user on the Web", in proceedings of the 5th Geographic Information Days (GIDAYS'07). 10-12 September 2007, Münster, Germany.

Contexte et Objectifs

Ce travail s'inscrit dans le contexte de la conception d'une carte sur mesure.

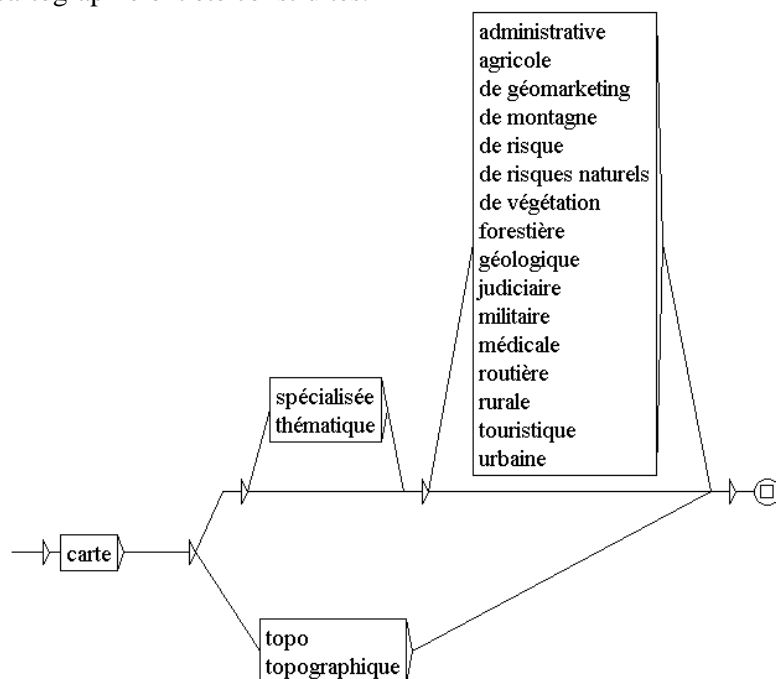
L'objectif est de construire les connaissances nécessaires à l'analyse et l'interprétation du besoin de l'utilisateur et à sa traduction en spécifications sur la carte à concevoir. Ces connaissances serviront de support à différentes tâches contribuant à la conception d'une carte sur mesure.

Résultats

Il faut établir des correspondances entre les mots des utilisateurs, exprimant leurs besoins et leurs goûts, et les concepts cartographiques pertinents permettant de construire la carte.

Dans ce but, plusieurs enquêtes ont été réalisées : la première, réalisée auprès de sujets expérimentés, a permis de collecter des descripteurs de cartes. Des échantillons cartographiques ont été élaborés puis qualifiés à l'aide de ces descripteurs par des sujets variés. Un corpus d'échantillons de cartes qualifiés a ainsi été construit. Une autre enquête s'est intéressée au ressenti des utilisateurs, exprimé en langage naturel. Des concepts variés (et le vocabulaire associé) permettant de décrire une carte ont ainsi été collectés et organisés. Ainsi un premier modèle de description formalisé de carte a été mis au point.

D'autre part, un corpus de textes de 150 000 mots provenant de source variées (manuels, livres, articles de journaux, commentaires de cartes, sites internet, etc) a été construit. Il rend compte des différents points de vue portés sur la carte par des publics variés utilisant des vocabulaires et des niveaux de langue différents. Il a été mis en forme pour permettre l'utilisation d'outils de traitement automatique de la langue. A partir de ce corpus, des ressources linguistiques (dictionnaires de mots simples et mots composés spécifiques des domaines, grammaires locales, etc.) spécifiques du domaine de la cartographie ont été construites.



La figure ci-contre montre sous la forme d'un automate, un extrait de la liste des termes qualifiant la destination de la carte.

Avec cet automate, on peut par exemple former les termes suivants:

carte thématique agricole
carte de risques naturels
carte topographique

Devant la multiplicité des termes retenus (environ un millier) et des relations (synonymie, antonymie, hyperonymie, hyponymie, relation tout/partie, etc.) qui les structurent, il est apparu nécessaire d'organiser ces termes dans une ontologie qui permettra de fixer le vocabulaire du domaine.

Pour en savoir plus...

Dominguès C., 2008, « Description de cartes géographiques ». 8èmes journées francophones Extraction et gestion des connaissances (EGC2008) – Atelier sur la modélisation utilisateur et la personnalisation d'interfaces web. Sophia-Antipolis. A paraître.

Dominguès C., 2007, « Du besoin de l'utilisateur à la spécification de sa carte », In Actes des 18èmes journées francophones d'ingénierie des connaissances (IC'2007), p.341-342 : Cépaduès-Éditions.

Contexte et Objectifs

Pour aider un utilisateur néophyte à construire une carte adaptée à ses goûts et ses besoins, une approche consiste à lui présenter des exemples de cartes dont il pourra utiliser intégralement la légende, ou dont il pourra s'inspirer pour construire sa propre légende.

Une enquête préalable a été menée (en 2005) auprès d'utilisateurs expérimentés afin de définir des propriétés pertinentes permettant de qualifier les propriétés visuelles d'une carte (*de couleurs chaudes, pastel, lumineuse, originale, sobre, ...*).

L'objectif est de construire une base de données d'échantillons originaux (par les choix de couleurs ou d'harmonies de couleurs) et cartographiquement corrects (les symboles utilisés dans les légendes respectent les relations d'association, différenciation et ordre qui lient les données).

Les propositions tiennent compte de la zone (à l'aide de deux paramètres *densité* et *relief*) que l'utilisateur veut représenter ainsi que des propriétés visuelles qu'il souhaite retrouver dans sa carte.

Résultats

Dans son stage, Pierre Philippe a d'abord créé une cinquantaine de palettes colorées. Celles-ci ont ensuite été appliquées à des données extraites de la BDTOPO® (zone de paramètre *densité* = *urbain* et *relief* = *mer*), à deux échelles 1/15 000 et 1/50 000, pour obtenir 104 échantillons qui ont enfin été évalués par rapport aux propriétés (*de couleurs chaudes, pastel, lumineuse, originale, sobre, ...*).



propriété : *originale*



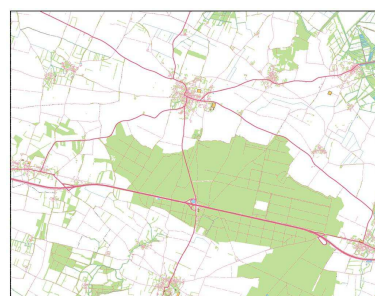
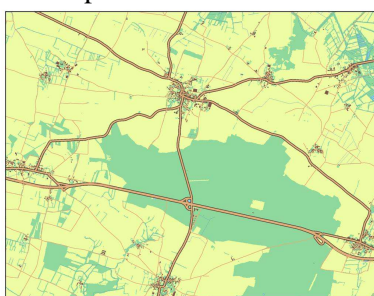
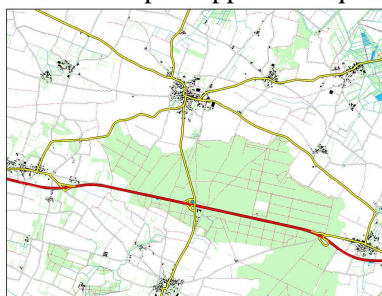
propriété : *de couleurs froides*



propriété : *de couleurs froides*

Présentation de trois échantillons de données BDTOPO® et de leurs propriétés

Le stage de Jérémy Renard consistait à étudier les légendes des cartes topographiques européennes. Des échantillons cartographiques ont été créés à partir d'extraits de la BDTOPO® auxquels ont été appliquées ces légendes. Un test utilisateur a été mené et analysé afin de préciser la perception de ces échantillons par rapport aux qualificatifs précédents et à l'efficacité de la visualisation des thèmes.



Présentation de trois échantillons de données BDTOPO® avec la symbolisation des cartes topographiques italienne, estonienne et finlandaise

Pour en savoir plus...

Dominguès C. (2007). « Conception et utilisation d'échantillons cartographiques pour l'aide à la conception de légendes ». In Bulletin de la recherche de l'IGN (à paraître)

Philippe P. (2006). « Définition d'un jeu d'échantillons cartographiques qualifiés », rapport de stage ENSG-IT3 et Master Carthagéo PRO, 41 pages, responsable de stage : C. Dominguès, laboratoire COGIT, mars 2006

Renard J. (2008). « Analyse et caractérisation de légendes européennes », rapport de stage ENSG-IT3 et Master Carthagéo PRO, 43 pages, responsable de stage : L. Jolivet, laboratoire COGIT, mars 2008

Contexte et Objectifs

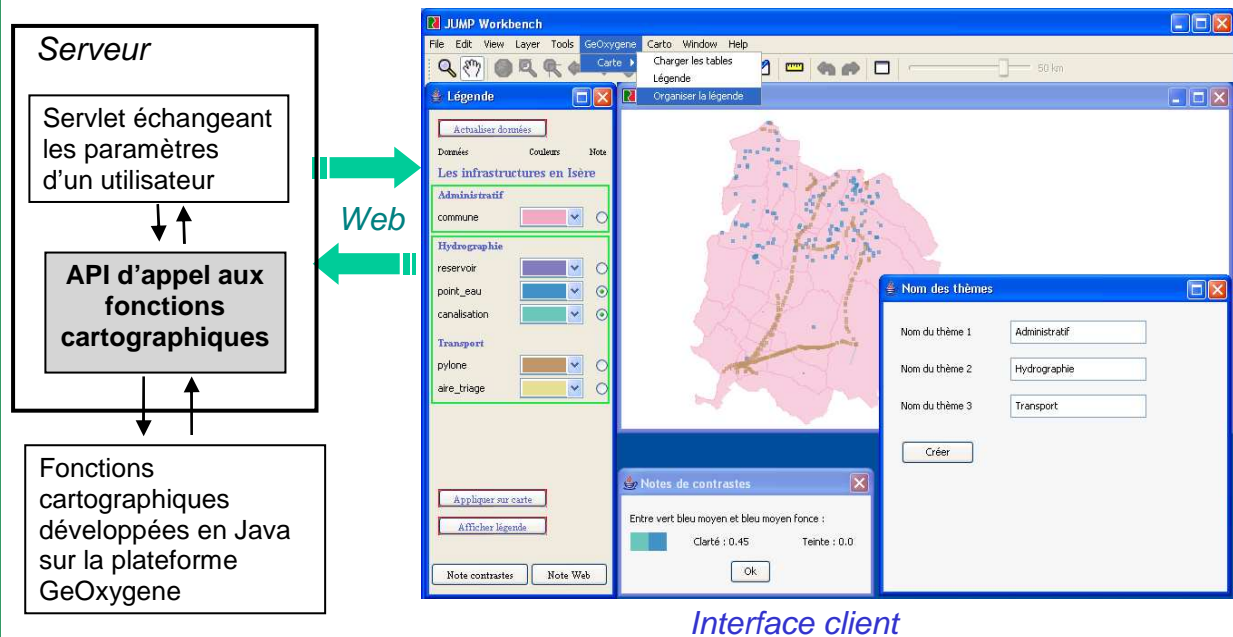
Ce travail se place dans le contexte de diffusion de données géographiques sur le Web. La personnalisation cartographique sur Internet concerne actuellement la sélection des couches à afficher et de la symbolisation à appliquer. Laurence Jolivet se concentre sur l'accessibilité de fonctions d'expertise cartographique qui améliorent la création de cartes sur mesure.

Le principal objectif est de développer une chaîne de services Web cartographiques qui prend en compte les étapes successives de réalisation de cartes. Ces services doivent s'appuyer sur les standards d'échange de données et d'informations sur Internet. Dans un premier temps, il s'agit d'identifier les méthodes existantes au COGIT qui se révèlent pertinentes lors d'une édition de cartes sur mesure. Par la suite, des fonctionnalités devront être développées afin de compléter ces méthodes et d'intégrer des principes cartographiques portant sur le choix des données et sur la définition de légende à partir de spécifications utilisateur.

Résultats

Une première application a été proposée à partir d'un client intégré au SIG OpenJUMP. La mise en œuvre s'est attachée à l'appel des traitements d'amélioration de légendes développées par Élodie Buard suite à la thèse d'Elisabeth Chesneau.

Un utilisateur a la possibilité de définir des spécifications de légende qui seront envoyées au serveur : les styles colorés, les thèmes, ainsi que les relations d'ordre, d'association ou de différence entre les données. Un service Web transmet ces paramètres et les applique aux fonctions spécifiques. Par exemple, une note de contraste coloré est calculée entre deux couleurs ou pour l'ensemble de la légende.



Application Web de création de légende et d'utilisation de fonctions cartographiques.

Le schéma ci-dessus expose le processus de communication entre le client et le serveur. Les points importants sont : l'interface graphique, l'utilisation des techniques d'échange sur le Web et l'appel aux fonctions cartographiques spécifiques favorisant la création d'une légende.

Pour en savoir plus...

Jolivet L., Buard E., Bucher B., Ruas A., 2007, « Amélioration de légende sur le Web », SAGEO, Clermont Ferrand, 18-20 juin 2007.

Jolivet, L., 2008, "On-demand map design based on user-oriented specifications". in proceedings of AutoCarto 2008 conference, Shepherdstown (WV), USA, 8-11 September 2008.

Contexte et Objectif

Le projet TerraNumerica du pôle de compétitivité Cap Digital vise à produire des données géographiques 3D extrêmement précises sur l'Ile de France. Il utilise essentiellement des méthodes de reconstruction appliquées à des images aériennes ou terrestres. Il étudie aussi la possibilité d'attacher à ce monde virtuel 'silencieux' des contenus qui ne sont pas visibles comme par exemple le parc locatif ou les fonctions des bâtiments et leurs horaires éventuels d'ouverture.

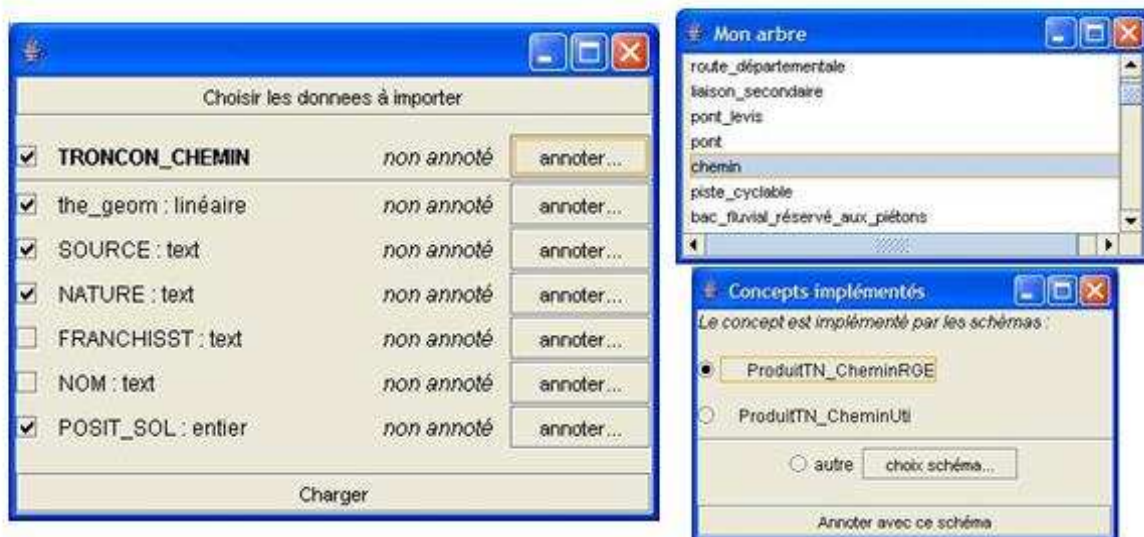
Les objectifs de ces travaux sont de fournir des services pour l'intégration de contenu utilisateur dans le monde 3D TerraNumerica. Les contenus visés peuvent être des flux dynamiques accessibles sur le Web, des bases de données ou un simple document.

Résultats

Une première phase a consisté à analyser le besoin, c'est-à-dire à voir dans quels contextes un fournisseur de contenu pourra enrichir le monde TerraNumerica avec ses informations. Cette analyse s'est faite au sein d'une équipe transversale au projet TerraNumerica et dédiée à l'ingénierie système. Dans cette équipe, B Bucher était responsable de la spécification des missions et capacités liées à l'enrichissement des données. Il a été nécessaire d'établir un glossaire de termes utilisés dans ce projet compte tenu de la diversité des domaines scientifiques (image, SIG, architecture).

L'interface de contenu proposée intègre des techniques de la géomatique et du Web collaboratif. Plusieurs niveaux d'intégration sont proposés :

- niveau 1 : le contenu utilisateur est référencé de la même façon que les autres informations TerraNumerica, avec le même modèle de métadonnées, dans des index communs.
- niveau 2 : le contenu est mis en cohérence avec les autres informations TerraNumerica, par exemple des services s'assurent que des abribus sont localisés à la surface des routes.
- niveau 3 : des liens sont créés entre des objets TerraNumerica et le contenu, par exemple entre la représentation d'un bâtiment et la liste des personnes célèbres ayant vécu dans ce bâtiment.
- niveau 4 : une représentation cartographique de ce contenu est spécifiée pour que les 'visiteurs' de TerraNumerica puissent le 'voir' dans ce monde virtuel.



Client permettant d'explorer les fonctionnalités disponibles sur GeOxygène, de sélectionner, paramétrer et invoquer une méthode et de comparer les résultats obtenus avec différents paramétrages

Pour en savoir plus...

Bucher B., WP 3.1.3.2 – D2, « Spécifications de services d'interface de contenu pour TerraNumerica », 2008

THEME 5 : ANALYSE DE L'ESPACE GEOGRAPHIQUE ET DE SES DYNAMIQUES

Par Jean-François Gleyze, responsable de l'action de recherche

CONTEXTE

En 1999, le laboratoire COGIT investit le domaine applicatif en consacrant une partie de ses travaux à l'étude de l'apport des bases de données géographiques dans l'analyse des risques.

Les recherches bibliographiques menées à cet effet montrent l'importance de l'information géographique et la pertinence de traitements spécifiques sur cette information pour comprendre et analyser les risques sur les territoires. De fait, les données géographiques ne servent pas seulement à décrire le support sur lequel se déroulent les phénomènes, mais elles contiennent également en puissance des informations sur les aspects spatiaux des risques.

La thèse soutenue en 2005 par Jean-François Gleyze a mis en évidence ce mécanisme dans l'étude de la vulnérabilité des réseaux de transport : dans un contexte de risques, il apparaît en effet que la structure sous-jacente des réseaux (déduite des seules bases de données géographiques) donne des informations sur les forces et les faiblesses de la trame des réseaux, et finalement sur la prédisposition de ceux-ci à subir des dysfonctionnements.

Au-delà de la thématique des risques et des réseaux de transport, le laboratoire COGIT propose d'apporter une expertise aux thématiciens dans leur utilisation des bases de données géographiques. En identifiant les besoins et les usages associés aux applications exploitant des données géographiques vectorielles, il s'agit de dépasser les simples considérations de géo-référencement, et d'extraire par des traitements spécifiques les informations implicites que ces données contiennent pour décrire et comprendre la dynamique spatiale des phénomènes territoriaux.

Cet apport s'appuie sur trois efforts de recherche :

- l'amélioration de la modélisation de l'information géographique ;
- le développement de connaissances sur la qualité de cette information ;
- l'enrichissement de l'information géographique par l'analyse spatiale des propriétés topologiques, géométriques, spatiales et sémantiques des objets décrits dans les bases de données associées.

OBJECTIFS

Ce thème de recherche a pour objectif principal de faciliter l'utilisation de données géographiques vectorielles à travers diverses applications utiles à la société.

Cette volonté d'ouverture se traduit par des objectifs particuliers, non seulement en termes de collaborations, de visibilité vers l'extérieur et de mission de service public, mais également en termes d'expertise scientifique sur l'information géographique au sein de l'IGN.

Avec les premiers résultats sur l'analyse des risques, le thème de recherche lié aux applications possède aujourd'hui l'expérience pour montrer le potentiel de l'information géographique vectorielle et engager des collaborations. Celles-ci doivent répondre aux missions de service public attendues pour l'IGN, et inscrites dans les orientations générales dictées par le Ministère de l'Écologie, de l'Énergie, du Développement Durable et de l'Aménagement du Territoire (MEEDDAT), et reprises dans le projet d'entreprise de l'IGN. Il s'agit à ce titre de tisser des collaborations qui permettent notamment de :

- mettre à disposition un *service public de visualisation et d'accès aux données géographiques*, avec l'élaboration de traitements d'analyse spatiale pour enrichir l'information géographique (directive européenne Inspire) ;
- favoriser la mise en place d'*études de risques, d'aménagement des réseaux ou de planification urbaine*, dans un souci de développement durable (texte de réforme du MEEDDAT, Direction Générale de la Prévention des Risques) ;
- participer à la *stratégie nationale pour la biodiversité*, par la *connaissance et la surveillance de l'évolution des écosystèmes et la préservation des populations d'espèce de la faune et de la flore sauvages* (propositions de recherche sur la biodiversité initiées par le Grenelle de l'Environnement).

Bien que tournés vers l'extérieur, ces orientations de recherche structurent en retour des objectifs scientifiques au sein du laboratoire, lesquels doivent se traduire à terme par un transfert de compétences vers l'IGN sous forme d'expertise et de développements.

Conformément à la ligne directrice du thème de recherche, les objectifs scientifiques doivent contribuer à l'amélioration des études thématiques exploitant des données géographiques vectorielles : en ce sens, ces objectifs se situent en amont des applications visées (phénomènes territoriaux, risques, dynamiques urbaines, réseaux, etc.).

1. Extraire de l'information implicite par analyse spatiale

Le premier de ces objectifs répond directement à la volonté de valoriser l'information géographique vectorielle. La pertinence de cette démarche a été démontrée dans le cas particulier de l'analyse des réseaux de transport dans un contexte de risques : le travail de Jean-François Gleyze a en effet montré que la connaissance *a priori* de la structure des réseaux de transport (telle qu'elle est décrite par les bases de données géographiques vectorielles) permet d'identifier les forces et les faiblesses structurelles des réseaux, et finalement de comprendre en quoi la géographie du réseau (*i.e.* les caractères topologiques, géométriques et spatiaux du réseau) prédispose aux perturbations dans un contexte d'usage. On notera aussi que l'automatisation du processus de généralisation a nécessité la conception de nombreuses méthodes d'analyse spatiale pour créer ou qualifier automatiquement des objets.

Sur la base de ces résultats, le laboratoire souhaite non seulement approfondir l'analyse des réseaux, mais également extraire des connaissances implicites sur tous les types d'objets géographiques vectoriels. Il s'agit en définitive de développer et de valoriser l'expertise du laboratoire en analyse spatiale, afin de permettre aux thématiciens de dépasser les simples considérations de géo-référencement usuellement associées aux données géographiques, et d'apprécier finalement dans quelle mesure celles-ci interviennent dans la spatialisation des phénomènes.

2. Mieux modéliser l'information géographique

L'ouverture du laboratoire sur des applications thématiques nécessite en amont d'améliorer et d'adapter la modélisation de l'information géographique, ainsi que de développer la connaissance de sa qualité.

Dans la perspective d'aider les thématiciens à mieux exploiter l'information géographique, il s'agit en premier lieu de proposer des méthodes et modèles adaptés aux applications visées, en particulier pour les données géographiques 3D, les données de terrain et les données relatives aux réseaux. Cette démarche doit notamment permettre de dépasser la relative 'neutralité' des formats de données et d'adapter celles-ci à leurs travaux applicatifs, en extrayant au besoin des informations dérivées pertinentes grâce à des traitements spécifiques, comme décrit plus haut.

Ce travail de modélisation renvoie par ailleurs aux questions de qualité des données et d'efficacité des traitements effectués sur ces données. Ce souci rejoint la volonté de l'IGN de garantir *la précision et l'efficacité des bases de données géographiques* (selon les termes du projet d'entreprise pour la période 2006-2015). Dans le contexte précis du thème de recherche lié aux applications, le laboratoire a pour objectif d'améliorer la qualité des analyses à support spatial en explicitant l'impact de la qualité des données vectorielles sur la précision géométrique de ces traitements.

3. Caractériser et simuler les évolutions du tissu urbain

Alors que les deux objectifs scientifiques énoncés précédemment dictent des orientations transversales, le laboratoire a défini un troisième objectif scientifique explicitement centré sur un type d'applications : l'étude des dynamiques territoriales.

Ce parti pris fait écho à l'importance accrue des études évolutives dans le domaine de l'écologie et de l'aménagement : qu'ils subissent des dégradations (risques) ou fassent l'objet de projets de développement (prospective), les territoires nécessitent en effet des analyses spécifiques pour comprendre la manière dont ils sont susceptibles d'évoluer, compte-tenu de l'existant (l'organisation de l'occupation du sol) et des modifications qu'ils sont susceptibles de subir (impact de risques, impacts d'aménagements, croissance urbaine). Ces analyses s'appuient sur une connaissance imbriquée de l'espace et de facteurs exogènes (du point de vue de l'expert en données géographiques),

pouvant concerner des domaines aussi variés que l'économie, la sociologie, le transport, etc. À la suite des travaux spécifiques menés sur les risques et les vulnérabilités, cet objectif de recherche a pour ambition de spécialiser l'expertise du laboratoire sur les méthodes d'analyse spatiale pour l'étude évolutive des territoires.

APPROCHES

L'approche choisie dans ce thème se décompose en trois axes, qui répondent collectivement à l'ensemble des objectifs énoncés plus haut (extraction d'information implicite par analyse spatiale, modélisation de l'information géographique, caractérisation et simulation des évolutions).

Analyse des réseaux

Le premier axe s'inscrit dans la suite des travaux sur la vulnérabilité des réseaux de transport : il s'intéresse à l'analyse spatiale des réseaux – non seulement les réseaux de transport, mais également toutes les structures réticulaires présentes explicitement ou implicitement dans les bases de données géographiques.

Au-delà des méthodes d'analyse élaborées dans le cadre précis de la vulnérabilité des réseaux de transport, les résultats développés dans la thèse de Jean-François Gleyze (formalisation, méthodologie, indicateurs) peuvent être sortis de leur contexte et généralisés, voire exportés vers d'autres domaines de recherches ou d'autres contextes d'étude.

L'analyse spatiale des réseaux intervient en effet dans de nombreuses applications thématiques, qu'il s'agisse d'applications spécifiquement ciblées sur des réseaux d'infrastructure (réseaux de transport, d'énergie, etc.) ou bien d'applications dans lesquelles les structures réticulaires interviennent dans la dynamique des phénomènes étudiés (diffusion, partition, hiérarchie spatiales, etc.). Conformément à la ligne directrice du thème de recherche, le laboratoire a pour ambition de montrer aux thématiciens que la structure spatiale des réseaux étudiés contient implicitement des connaissances pouvant les aider dans leur application. À ce titre, l'approche retenue pour cet axe se décompose en deux parties :

1. formaliser l'analyse structurelle des réseaux et élaborer des indicateurs quantitatifs sur les potentialités des réseaux
2. intégrer l'ensemble des méthodes et des indicateurs élaborés au sein d'un outil d'analyse structurelle, dont l'objet est d'organiser, explorer, représenter et visualiser les informations structurelles relatives aux réseaux.

Le deuxième axe de recherche concerne l'analyse spatio-temporelle de l'espace : Notre recherche s'inscrit dans la continuité de recherches réalisées précédemment au COGIT, à savoir le travail de thèse d'Annabelle Boffet sur la caractérisation des zones urbaines, des quartiers et des îlots (2001) et le travail de thèse de Patricia Bordin sur le suivi de phénomènes géographiques avec un système d'information géographique (2006). Cet axe est particulièrement mis en valeur depuis que le laboratoire coordonne le projet ANR GeOpenSim (2007-2010), dont l'objectif de concevoir une plateforme Open Source dédiée à l'analyse des tissus urbains. L'approche retenue repose sur un système multi agents et une représentation vectorielle et multi niveaux des objets géographiques, dans le but d'assister les études d'implémentation de programmes architecturaux ou de planification, d'enrichir les recherches sur la morphogénèse des systèmes urbains et surtout d'améliorer la compréhension des mécanismes urbains évolutifs. À terme, un tel outil doit permettre d'identifier l'émergence d'organisations spatiales et de qualifier les temporalités associées à ces évolutions.

Le troisième axe concerne l'étude des données géographiques vectorielles et les questions de modélisation et de qualité qui y sont associées.

Sur la base des travaux de recherche d'Olivier Bonin et de Benoit Poupeau sur la simplification des objets 3D et l'étude des interactions spatiales entre objets 3D, le laboratoire souhaite développer plus avant son expertise sur ce type de données, afin de répondre aux demandes croissantes des thématiciens utilisant les représentations 3D dans la compréhension des risques et des phénomènes naturels. À cet effet, le laboratoire oriente ses efforts sur l'élaboration d'un modèle tridimensionnel de l'espace géographique, dont les méthodes permettraient :

- l'analyse de données géographiques 3D (par exemple dans un contexte d'étude de risques) ;

- l'intégration de données géographiques 3D (par exemple dans la mise en commun de plusieurs sources de données géographiques, de qualité hétérogène) ;
- et enfin l'échange de données géographiques 3D (notamment dans le cadre du respect des normes ISO et OGC).

Le laboratoire souhaite également disposer d'une expertise sur la *qualité des données vectorielles* et l'incidence de celle-ci sur la qualité des études menées par les thématiciens. Ce choix s'appuie sur des recherches spécifiques menées antérieurement par le laboratoire COGIT, sur les imprécisions géométriques (travaux de François Vauglin), sur le calcul d'itinéraires (travaux d'Olivier Bonin) et sur l'appariement de données (travaux d'Ana-Maria Olteanu). L'approche retenue aujourd'hui a pour objet d'établir les rapports entre la qualité en entrée (qualité des données vectorielles) et la qualité en sortie (qualité des résultats) des traitements opérés par les utilisateurs de l'information géographique. À ce titre, il ne s'agit pas tant de quantifier précisément la qualité des données et des traitements (d'autant que les données en entrée peuvent être peu qualifiées), mais de proposer à l'utilisateur des indications sur les niveaux attendus de qualité, dans la perspective d'être intégrés à des processus décisionnels (par exemple dans le cadre d'études de risques).

PRINCIPAUX RESULTATS ET PERSPECTIVES

Les travaux de ce thème de recherche s'appuient sur des applications, aussi les résultats obtenus s'inscrivent-ils avant tout dans le contexte thématique de ces applications.

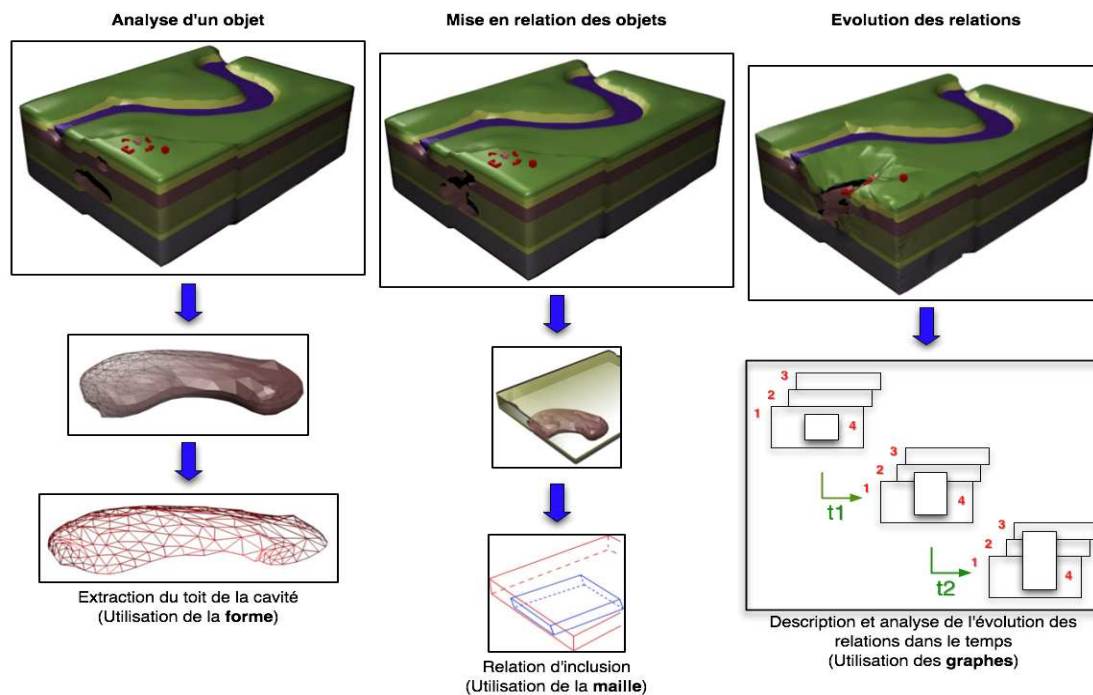
- Jean-François Gleyze a développé des indicateurs quantitatifs pour évaluer les potentialités du réseau dans un contexte de risques : en ce sens, ces indicateurs donnent des informations sur les capacités de réaction ou d'adaptation des réseaux de transport dans l'éventualité d'endommagements localisés (capacité de remplacement offertes par les composantes du réseau, aptitude à être remplacées, évaluation des alternatives offertes par le réseau pour chaque relation origine-destination, etc.).

La partie applicative de ces recherches est prolongée par les travaux de thèse d'Éric Mermet sur la conception d'un outil d'exploration et de visualisation des réseaux. Les premiers résultats proposent des développements sur les changements de représentation et les requêtes sélectives dans le calcul et la visualisation des indicateurs, et s'orientent vers la conception d'un langage graphique accessible.

- Adrien Paget a initié l'analyse systématique des réseaux hydrographiques grâce à l'implémentation d'indicateurs permettant l'identification de motifs récurrents et la segmentation des réseaux en plusieurs catégories (réseaux parallèles, dendritiques, etc.), conformément aux typologies établies par les géomorphologues.

- Olivier Bonin et Benoit Poupeau ont mis au point des méthodes de modélisation et de simplification des objets 3D pour analyser les interactions spatiales de voisinage de bâtiments et de couches géologiques dans le cadre de travaux sur les effondrements de cavités dans le bassin minier de Lens (voir figure suivante).

- Patricia Bordin a mis au point des méthodes de suivi de phénomènes géographiques à l'aide de systèmes d'information géographique (SIG). Elle montre que les SIG offrent des solutions techniques et spécifiques, requérant généralement des développements informatiques à façon, permettant d'intégrer des données de sources multiples pour créer des indicateurs relevant de niveaux différents et de domaines de caractérisation variés.



Analyse des objets – Mise en relation des objets – Evolution des relations' (Benoit Poupeau et d'Olivier Bonin)

Ces résultats s'inscrivent dans une large palette de domaines applicatifs et font appel à des techniques spécifiques de l'information géographique (SIG, indicateurs quantitatifs, représentations 3D, topologie, etc.). Pour autant, leurs apports convergent en regard des objectifs scientifiques énoncés pour ce thème de recherche. Les travaux développés dans ce thème proposent en effet des cadres d'étude, des méthodes et des traitements d'analyse spatiale innovants et adaptables à d'autres domaines d'études, ouvrant autant de perspectives de recherche à moyen terme.

- Les travaux sur les réseaux ont débouché sur un ensemble unifié de méthodes et d'indicateurs fondés sur la théorie des graphes. Cet ensemble repose sur un cadre d'étude plaçant la structure des réseaux au centre de l'analyse. En premier lieu, cette approche structurale offre une base théorique permettant de tester et d'élaborer de nouveaux indicateurs structurels fondés sur l'observation des chemins (les plus courts chemins et les chemins alternatifs). Les résultats les plus récents ont notamment permis d'identifier les hiérarchies entre composantes induites par la structure des chemins (indicateur de dépendance relationnelle), de nuancer la notion d'accessibilité en tenant compte des chemins alternatifs (indicateur du nombre de chemins équivalents au plus court chemins) et enfin de mesurer pour tout indicateur l'apport net du réseau dans le phénomène sous-jacent, en enlevant le biais spatial. Cette approche structurale ouvre des perspectives sur l'étude des structures en réseau pouvant être extraites des bases de données géographiques (réseaux d'infrastructure, mais également réseaux de voisinages ou réseaux hiérarchiques construits sur des classes d'objets géographiques).
- Les travaux sur les réseaux hydrographiques ont abouti à l'élaboration d'indicateurs géométriques pour l'identification de réseaux caractéristiques, en accord avec la typologie admise par les géomorphologues. Au-delà de l'application immédiate de ce travail pour l'analyse de l'hydrographie, ces indicateurs sont susceptibles de fournir des informations sur les types de relief sous-jacents. Ces informations peuvent aider à la définition de zones géographiques homogènes du relief (les Modèles Numériques de Terrain) et apporter des connaissances sur la qualité de ces données : cette perspective constitue un tremplin pour les travaux à venir dans l'axe de recherche dédié à la qualité.
- Les travaux sur le SIG 3D ont abouti à l'élaboration d'algorithmes de simplification des bâtiments 3D. Les résultats des algorithmes ont pour but de respecter au mieux la forme des bâtiments initiaux tout en proposant différents niveaux de simplification. Ces résultats sont utiles pour représenter de manière plus légère les éléments en arrière plan. De plus, les formes simplifiées permettent

d'optimiser le temps de calcul pour certaines opérations complexes, tout en obtenant des résultats satisfaisants. Un algorithme mesurant un indicateur de ressemblance entre objets a également été mis en place pour quantifier la ressemblance entre l'objet initial et l'objet simplifié. Cet algorithme pourrait être utilisé pour déterminer l'algorithme de simplification le plus judicieux à appliquer.

- Enfin, les travaux sur les dynamiques spatiales ont permis de définir les premières briques d'un cadre formel innovant permettant l'analyse et le suivi des évolutions des tissus urbains. Ceci s'est notamment traduit par le développement d'outils de base de la plate-forme Open Source GeOpenSim, parmi lesquels figurent des algorithmes d'analyse spatiale, des schémas de données géographiques et les fondements du système multi-agents au cœur du système. La simulation et les mécanismes d'apprentissage permettront au système de révéler l'émergence d'organisations spatiales, mais aussi de qualifier et d'enrichir les connaissances associées à ces évolutions et à leurs temporalités. Ces travaux élargissent nos réflexions sur les données historiques et sur la gestion du temps dans les systèmes d'information géographique.

À terme, et en relation avec d'autres spécialistes, une telle approche pourrait intégrer des agents sociaux – individus, organisations – afin d'étudier les rétroactions des évolutions urbaines sur la société et inversement des dynamiques sociales sur la morphologie urbaine. Une telle approche écosystémique permettrait d'apporter une modélisation innovante pour l'analyse des interactions entre une société et son environnement.

Contexte et objectifs

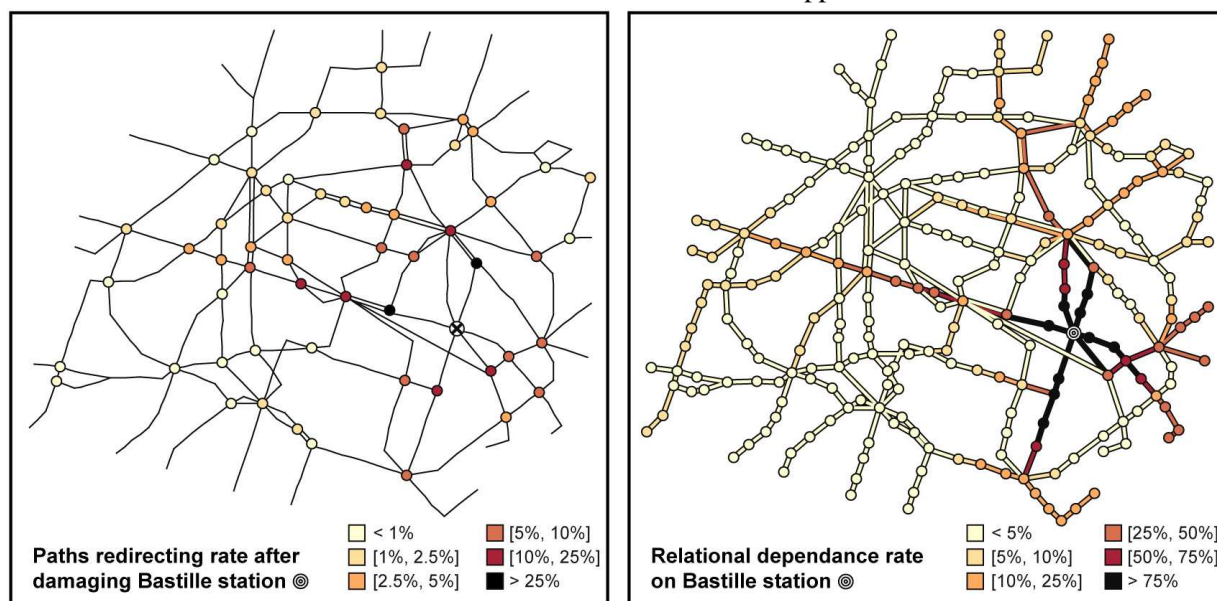
L'analyse structurelle des réseaux de transport s'inscrit dans la connaissance des risques et l'extraction de connaissances sur les structures en réseau. Ce croisement de problématiques a conduit Jean-François Gleyze à travailler sur la vulnérabilité des réseaux de transport dans un contexte de risques (thèse soutenue en 2005). Le cadre d'étude, les méthodes et les indicateurs développés ouvrent des perspectives sur l'étude structurelle des réseaux et la mise au point d'indicateurs pour comprendre les potentialités des réseaux : leurs forces et leurs faiblesses, leur capacité de réaction ou d'adaptation dans des contextes évolutifs.

Approches et résultats

L'analyse structurelle d'un réseau consiste à se focaliser sur sa trame, décrite exclusivement par la base de données géographiques. Tes traitements associés ont donc pour objet d'extraire des connaissances implicites sur ce réseau, à partir de ses caractéristiques topologiques, spatiales, géométriques et métriques. Cette approche apporte des informations sur les potentialités du réseau pour éclairer son fonctionnement contextuel.

Ce cadre d'étude permet d'élaborer des indicateurs, représentatifs des caractéristiques des chemins offerts par le réseau, non seulement les plus courts chemins, mais également toute classe de chemins alternatifs : cette éventualité permet d'intégrer le caractère évolutif du réseau.

Les figures ci-dessous illustrent deux indicateurs développés dans ce cadre : un indicateur pour localiser et quantifier les changements d'itinéraires dans l'éventualité d'un endommagement local du réseau (figure de gauche), et un indicateur pour mesurer les dépendances relationnelles des composantes du réseau vis-à-vis d'une composante donnée (figure de droite). D'autres indicateurs sur le rayon d'action des composantes au sein du réseau ou sur l'évaluation de la contribution des chemins alternatifs à l'accessibilité au sein du réseau sont en cours de développement.



Informations sur la structure des plus courts chemins transitant par la station Bastille sur le réseau de métro parisien – taux de redirection dans les stations à correspondance dans l'éventualité d'une indisponibilité de la station Bastille (à gauche) et taux de dépendance relationnelle des composantes du réseau à la station Bastille (à droite) (d'après : Gleyze, ECTQG 2007)

Pour en savoir plus...

Gleyze J.-F., 2007, « Effets spatiaux et effets réseaux dans l'évaluation d'indicateurs sur les nœuds d'un réseau d'infrastructures. » CyberGeo, <<http://www.cybergeo.eu/index5532.html>>, n°370, 24 p.

Gleyze J.-F., Reghezza M., 2007, « La vulnérabilité structurelle comme outil de compréhension des mécanismes d'endommagement ». Géocarrefour, Vol. 82, n°1-2, numéro spécial : « Le risque, de la recherche à la gestion territorialisée », p. 17-26

Contexte et objectifs

Thèse de doctorat en Sciences de l'Information Géographique de l'Université Marne-La-Vallée, dirigée par Anne Ruas, et co-encadrée par Jean-François Gleyze, démarrée fin 2007

L'objectif de cette thèse est de mettre au point des méthodes d'exploration de réseaux grâce à un outil de visualisation interactif. Il s'agit de mettre en évidence les potentialités offertes par un réseau pour mettre en relation les lieux sur le territoire.

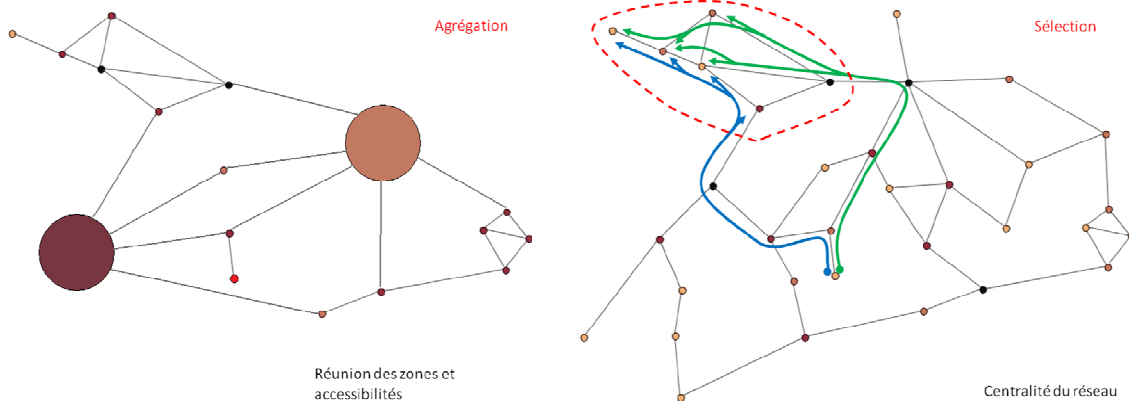
Il faudra ainsi 1/ concevoir des indicateurs d'analyse pertinents, 2/ concevoir des méthodes permettant d'agréger ces indicateurs puis 3/ proposer un langage graphique permettant la manipulation aisée des indicateurs et des agrégats dans le but d'explorer le réseau et les informations calculées qui s'y rattachent.

Premiers résultats

Le travail de première année a été centré sur les méthodes d'exploration par changement de représentation et de requêtes sélectives des relations OD (origine-destination). Ces méthodes se heurtent à la grande taille des données, à la combinatoire des graphes, et à la complexité du calcul des indicateurs (les algorithmes associés étant souvent NP-complets).

Les changements de représentation expérimentés cette année sont produits par des agrégations locales au sein du réseau (figure de gauche). L'agrégation d'un sous-réseau en un nœud unique permet de cerner certaines zones du réseau et d'obtenir des informations agrégées sur ces zones. Cette méthode exige par ailleurs de déterminer des règles pour résumer les indicateurs sur des zones agrégées, étant entendu que ces règles dépendent de la structure des zones et du type d'indicateur.

Les requêtes sélectives qui ont été testées sur les relations OD consistent à déterminer manuellement la collection des relations origine-destination sur lesquelles le calcul des indicateurs sera restreint (figure de droite) : cette méthode permet de comprendre comment les différentes relations OD contribuent aux valeurs des indicateurs, et finalement de cerner les singularités locales du réseau. Pour autant, cette méthode devra être approfondie par la mise en place de requêtes sélectives suggérées ou déduites de traitements.



exploration par changement de représentation

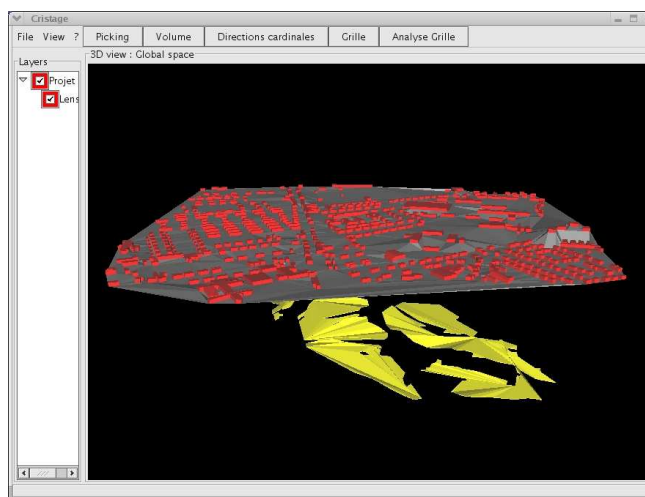
exploration par requêtes sélectives sur les relations OD

L'étape suivante va consister en un enrichissement des indicateurs (notamment en intégrant les chemins alternatifs), en la conception de méthodes d'agrégation dynamique et en l'utilisation d'un langage graphique permettant la construction des requêtes.

Contexte

Olivier Bonin a mené avec Benoît Poupeau (Paris-Est) une étude morphologique du bassin minier de Lens, dans le but de mettre en évidence des mouvements locaux dus aux exploitations minières et d'éventuels mouvements plus généraux, à l'aide d'un modèle numérique de terrain de la zone reflétant le terrain post 1959 (exploitations minières en cours), et un modèle numérique de terrain post exploitations minières.

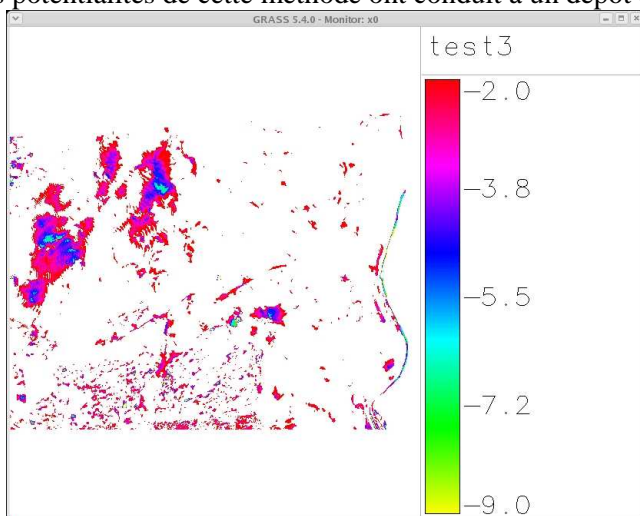
La figure suivante illustre un quartier de Lens, le modèle numérique de terrain et une veine de charbon exploitée, entièrement modélisés en 3D dans Cristage (voir travaux de Benoît Poupeau). Le système permet de répondre à des requêtes spatiales : Que trouve-t-on en dessous de telle localisation ? Quel est le volume des veines exploitées en telle année ? Telle zone et telles veines sont-elles en relation ? Etc.



Extrait d'un quartier de Lens et d'une veine dans Cristage. Données minières : Charbonnage de France.

Approches et Résultats

Olivier Bonin a mis au point une méthode de comparaison de modèles numériques de terrain d'actualité différentes, en collaboration avec Benoît Deffontaines (IFG-Paris-Est). Cette méthode de comparaison permet de suivre les déformations cumulées de la surface topographique dans le temps. Elle a été testée sur des zones d'anciennes exploitations minières (bassin houiller du Nord-Pas-de-Calais, potasses d'Alsace ; mines de Gardanne, exploitations de gaz et de pétrole de la région de Lacq). Les potentialités de cette méthode ont conduit à un dépôt d'un brevet IGN/Paris-Est (en cours).



Carte des écarts significatifs entre MNT récent et MNT historique (une centaine d'années d'écart) sur la feuille de Mulhouse : les écarts négatifs importants au nord-ouest correspondent aux affaissements liés à l'exploitation souterraine des Potasses d'Alsace. L'aspect patchwork au sud correspond à la différence de résolution et l'évolution du MNT dans la ville de Mulhouse entre la BDAlti historique et la BDTopo.

Pour en savoir plus ...

Bonin O. ; Deffontaines B., Poupeau B. ; Gueguen Y et B. Fruneau, 2006, « Combinig ground displacement information with a 3D geographical information system to perform post-mining diagnosis », accepté à la conférence de l'ISPRS Commission VIII, Haïfa, Israel.

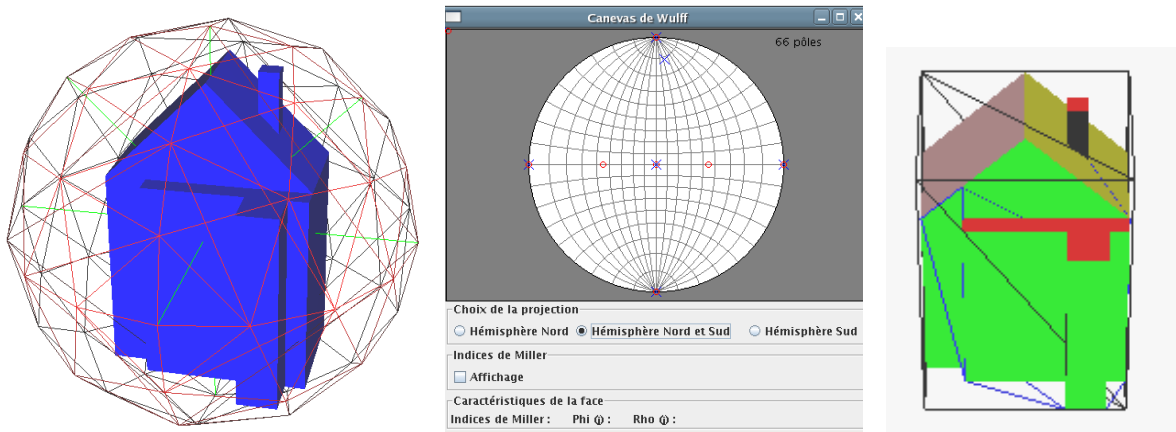
Contexte et objectifs

Thèse de doctorat en Sciences de l'Information Géographique de l'Université Paris-Est, soutenue en septembre 2008, dirigée par Anne Ruas, et co-encadrée par Olivier Bonin

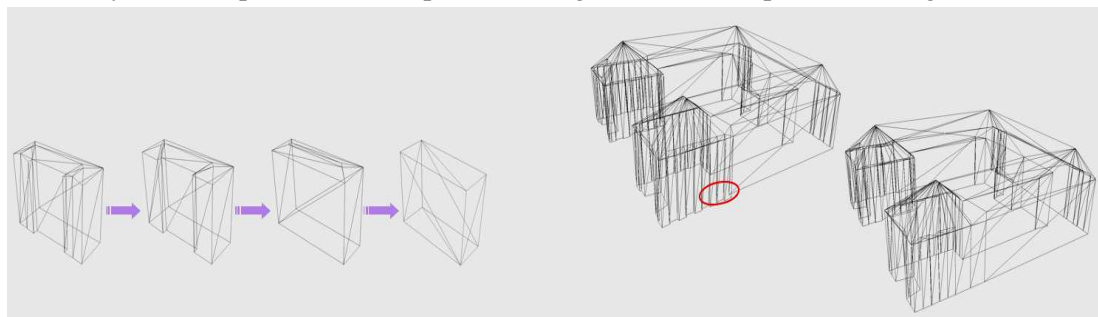
L'objectif de cette thèse est de mettre au point un modèle et des méthodes pour faciliter l'analyse de données 3D dans un contexte de risque. Le modèle proposé doit permettre des analyses intra et inter-objets notamment par la représentation des positionnements relatifs significatifs pour les phénomènes d'effondrement et d'écoulement.

Résultats

Les recherches de Benoît Poupeau sur l'analyse intra-objet pour décrire et manipuler les objets géographiques ont abouti 1/ sur la notion de structure, qui décrit la manière dont les primitives géométriques s'agencent entre elles, et 2/ sur l'intégration de principes issus de la cristallographie géométrique. Complémentaire des modèles topologiques, l'analyse de la structure repose sur l'extraction de la symétrie (axes, plans centre) des objets à l'aide d'une projection stéréographique (figure ci-dessous). Elle facilite la description sémantique, la classification en fonction de la symétrie, l'extraction de certaines parties caractéristiques comme le toit d'un bâtiment ou d'une cavité et la manipulation des objets géographiques 3D.



Analyse des symétries d'un bâtiment, construction de maille englobante et reconnaissance des toits
L'étude des symétries a permis la conception d'un algorithme de simplification (régularisation)



Simplification 3D d'une cavité (à gauche) et d'un bâtiment (à droite).

Au delà de la modélisation, la thèse propose des pistes intéressantes sur les espaces proximaux et les graphes temporels pour le suivi de l'évolution des relations entre objets (dont les effondrements). Une partie du code développé sous le nom de Cristage sera déposé en OpenSource (page suivante).

Pour en savoir plus

Poupeau B., 2008, « Analyses et Requêtes de données géographiques 3D – Contributions de la cristallographie géométrique ». Thèse de doctorat Paris-Est, ICMS, soutenue le 16 Septembre 2008.

Poupeau B. et Ruas A., 2007, "A crystallographic approach to simplify 3D building". In proceedings of International Conference of Cartography, Moscou, Russie

Contexte et objectifs

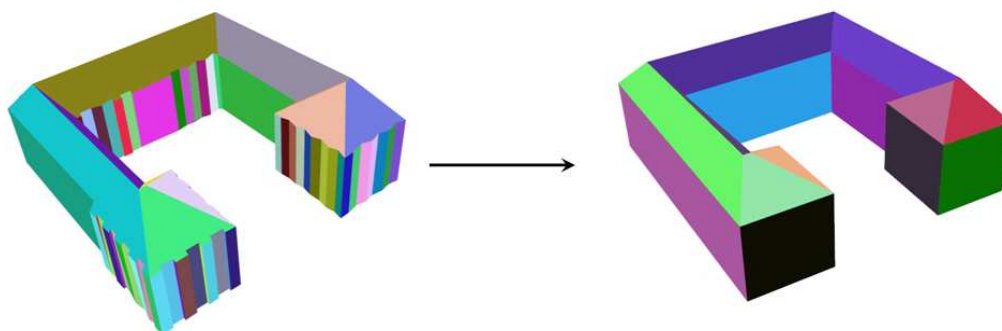
A la suite du travail de thèse de Benoit Poupeau (page précédente), l'objectif de ce travail est de réaliser un module 3D sur la plate-forme GeOxygène et de le déposer en OpenSource. Par rapport au code existant, il s'agit de se rapprocher des normes ISO et de la philosophie de GeOxygène. Ce module comportera des algorithmes de manipulation, de visualisation et de simplification d'objets 3D.

Résultats

La structure de modélisation géométrique choisie est la représentation par frontières (BRep) qui consiste à représenter un solide comme étant composé d'un ensemble de faces délimitées par des arcs reliant les sommets de la géométrie. Différentes bibliothèques libres ont été utilisées afin d'introduire les fonctionnalités fondamentales d'un SIG 3D :

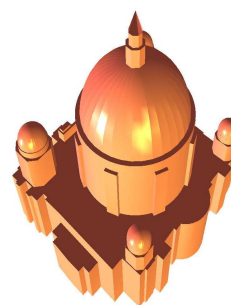
- l'affichage de géométries 3D ;
- la tétraèdrisation d'objets 3D ;
- le calcul d'opérations booléennes en 3D.

Un algorithme de simplification de bâtiment 3D a été développé sur cette plate-forme. Cet algorithme opère à partir de la trace au sol des murs du bâtiment. La simplification a lieu par fusion des zones tampon de ces traces, puis par reconstruction 3D, face par face, à partir du résultat de cette fusion. Le toit est ensuite adapté sur la forme du nouveau solide.



Simplification d'un bâtiment 3D (Bati3D - Hôtel de ville d'Amiens)

Par ailleurs, des développements spécifiques ont été effectués permettant l'intégration de données CityGML et de données du projet Bati3D de l'IGN.



Jeu de données au format CityGML sur Ettenheim (Allemagne) à gauche et le Berliner Dom à droite.

Un dépôt Open-Source contenant l'outil de visualisation 3D et l'algorithme de simplification devrait avoir lieu début 2009.

Pour en savoir plus ...

Velten A, 2007, « Implémentation d'un algorithme de généralisation automatique de bâtiment 3D », Rapport de stage, septembre 2007.

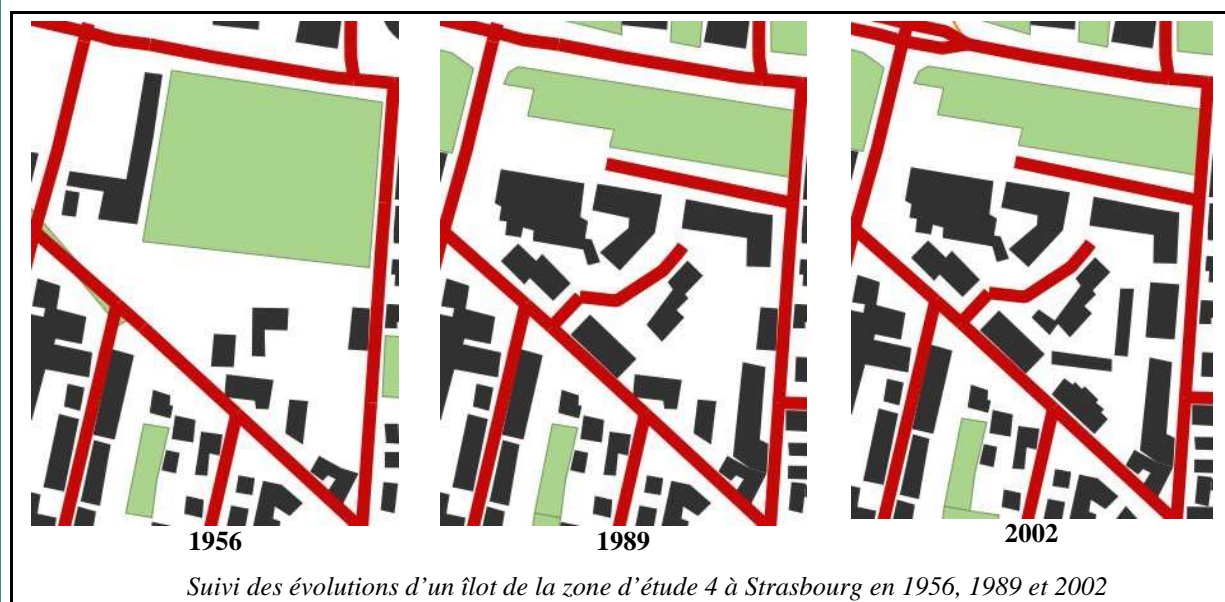
Marcé L, 2007, « Développement d'une IHM 3D pour le moteur de SIG3D Cristage », Rapport de stage, septembre 2007

Contexte

Ces recherches se font dans le cadre du projet ANR GeOpenSim sur la conception d'une plate-forme Open Source dédiée à l'analyse des tissus urbains et de leurs évolutions. Ce projet, commencé en décembre 2007 et d'une durée de 3 ans, est coordonné par Anne Ruas et fait l'objet d'une collaboration avec le Laboratoire Image et Ville (LIV) et le Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (LSIIT) à Strasbourg, ainsi qu'avec Annabelle Boffet-Mas du Centre d'Études pour le Développement des Territoires et l'Environnement (CEDETE) à Orléans.

Approches et Résultats

L'approche du projet GeOpenSim repose sur une analyse morphologique et topologique des évolutions subies par les objets géographiques entre des versions successives de bases de données géographiques. Cette analyse diachronique des villes s'appuie sur une représentation vectorielle et multi-niveaux de l'information géographique et permet d'étudier les dynamiques urbaines unitaires, mais aussi les mécanismes de déclenchement de celles-ci. La représentation de connaissances sur ces mécanismes sous la forme de règles permet leur intégration au sein d'un système multi-agents qui effectue la simulation des évolutions du système dans le temps.



Les premiers résultats sont composés 1/ d'une première typologie des opérations unitaires d'évolution des objets géographiques, 2/ d'un schéma de données permettant de porter à les différentes temporalités des bases de données historiques et b- la simulation à base d'agents vectoriels 3/ du développement d'algorithmes d'analyse spatiale et des mécanismes de simulation du système multi-agents.

En 2008, les partenaires du projet apportent leur contribution sur l'analyse géographique des évolutions, la constitution de BD Historiques et la constitution de bases d'exemples pour la classification automatique des îlots.

En 2009 les premiers patterns d'évolution devraient être identifiés (entre autre par apprentissage) et les premières fonctions de peuplement devraient être développées.

Pour en savoir plus ...
<http://geopensim.ign.fr/>

TROISIEME PARTIE

**PUBLICATIONS ET COMMUNICATIONS
2006-2008**

THESES DE DOCTORAT

2008

Gaffuri Julien, 2008. *Généralisation automatique pour la prise en compte de thèmes champs: le modèle GAEL*, thèse de doctorat en sciences de l'information géographique, université Paris-Est, Laboratoire COGIT.

Olteanu-Raimond Ana-Maria, 2008. *Fusion de connaissances imparfaites pour l'appariement de données géographiques : proposition d'une approche s'appuyant sur la théorie des fonctions de croyance* Thèse de Doctorat de l'Université Paris Est, soutenue le 24 octobre 2008.

Poupeau Benoit, 2008, *Analyses et Requêtes de données géographiques 3D - Contributions de la cristallographie géométrique*. Thèse de doctorat de l'Université de Paris-Est / Laboratoire COGIT. Soutenue le 16 Septembre 2008.

Pierkot Christelle, 2008. *Gestion de la Mise à Jour de Données Géographiques Répliquées*. Thèse de Doctorat de l'université Paul Sabatier Toulouse III, soutenue le 2 juillet 2008.

Taillandier Patrick, 2008, *Révision automatique des connaissances guidant l'exploration informée d'arbres d'états. Application au contexte de la généralisation de données géographiques* Thèse de doctorat de l'université Paris-Est / Laboratoire COGIT, soutenue le 2 décembre 2008

2007

Balley Sandrine, 2007, *Aide à la restructuration de données géographiques sur le Web - Vers la diffusion à la carte d'information géographique*. Thèse en informatique de l'Université de Paris-Est / Laboratoire COGIT, septembre 2007

2006

Abd El Kader Y., 2006, *Conception et exploitation d'une base de métadonnées de traitements informatiques, représentation opérationnelle des connaissances d'expert - Application au domaine géographique*, Thèse en informatique de l'Université de Caen, Saint Mandé, juillet 2006

Bordin Patricia, 2006, *Méthode d'observation multi-niveaux pour le suivi de phénomènes géographiques dans un SIG*, Thèse de doctorat en Sciences de l'Information Géographique, Université de Paris-Est Marne-La-Vallée, soutenue le 12 décembre 2006

Chesneau Elisabeth, 2006, *Modèle d'amélioration automatique des contrastes de couleurs en cartographie - Application aux cartes de risque*, Thèse de doctorat, Université de Paris-Est Marne La Vallée, soutenue le 23 mai 2006

ACTES DE CONFÉRENCES

Ruas A. & C. Gold, 2008, *Headway in Spatial Data Handling*, Lectures notes in Geoinformation and Cartography, Springer, Spatial Data Handling notes (SDH'08, Montpellier)

EDITION D'OUVRAGES

Mackaness W., **Ruas A.**, Sarjakoski L.T., 2007, *Generalisation of geographic information: cartographic modelling and applications*, Elsevier eds

CHAPITRES D'OUVRAGE

Mackaness W., **Ruas A.**, 2007, Evaluation in the Map generalisation Process, Dans *The Generalisation of Geographic Information : Models and Applications*, sous la direction de W. Mackaness, A. Ruas et T. Sarjakoski. Elsevier Chapter 5

Mustière S., Van Smaalen J., 2007, Database Requirements for Generalisation and Multiple Representations. Dans *The Generalisation of Geographic Information : Models and Applications*, sous la direction de W. Mackaness, A. Ruas et T. Sarjakoski. Elsevier, Chapitre 6.

Olteanu A.-M. 2008. A Multi criteria fusion approach for geographical data matching. *Quality Aspects in Spatial Data Mining*. Taylor and Francis (éd.), ISBN: 9781420069266, pp. 45-56.

Ruas A., Duchêne C., 2007, A prototype generalisation system based on multi-agent system paradigm, Dans *The Generalisation of Geographic Information : Models and Applications*, sous la direction de W. Mackaness, A. Ruas et T. Sarjakoski. Chapter 14

REVUES A COMITE DE LECTURE

2008

Christophe S., 2008, *Legend Design on The Web: Creating Accurate Styles*. *International Journal of Spatial Data Infrastructures Research (IJSDIR) Special Issue 25 pages*, Vol.3 (2008), Guest Editors: Florian Probst and Carsten Kessler

Hubert, Frédéric. 2008 *CartAble, système d'aide au paramétrage de traitements géographiques complexes*. *Revue internationale de géomatique*, Vol. 18-n°1/2008 pp.41-65.

Mustière S., Devogele D. 2008. *Matching networks with different levels of detail*. GeoInformatica, vol.12, n.4, 12/2008, pp.435-453.

Paget A., Perret J., Gleyze J.-F., 2008, *L'analyse des graphes hydrographiques pour la caractérisation du relief*. Article à paraître dans la revue en ligne Physio-Géo, 18 p.

Sheeren, D., Mustière, S. and Zucker, J.-D., 2008, *A data-mining approach for assessing consistency between multiple representations in spatial databases*, International Journal of Geographical Information Science, Publication en ligne en 2008, publication dans la version papier du journal en 2009.

2007

Gleyze J.-F., 2007, *Effets spatiaux et effets réseaux dans l'évaluation d'indicateurs sur les noeuds d'un réseau d'infrastructure*, CyberGeo, n°370, 24p

Gleyze J.-F., Reghezza M., 2007, *La vulnérabilité structurelle comme outil de compréhension des mécanismes d'endommagement*, Géocarrefour, Vol.81, numéro spécial "Le risque, de la recherche à la gestion territorialisée", 14p.

Purves R.S., Clough P., Jones C.B., Arampatzis A., Bucher B., Finch D., Fu G., Joho H., Khirini A.S., Vaid S., Yang B., 2007, *The design and implementation of SPIRIT: a spatially-aware search engine for information retrieval on the Internet*, International Journal Geographic Information Systems (IJGIS), Vol21(7), January 2007, pp. 717-745

2006

Bucher B., 2006, *Décrire formellement les processus de manipulation de données géographiques*, revue Ingénierie des systèmes d'Information (ISI), Systèmes d'Information spécialisés, Patrick Bosc (Ed), vol 11, n°1

Chesneau E., 2006, *Propositions méthodologiques pour l'amélioration automatique des contrastes de couleur - Application aux cartes de risque*, Cybergeo, n°360, 14 novembre 2006

CONFÉRENCES A (Sélection sur articles long, comité de lecture, publication)

2008

Abadie N. et Mustière S. 2008. *Constitution d'une taxonomie géographique à partir des spécifications de bases de données*. Actes de SAGEO'2008, Montpellier, 26 juin 2008.

Duchêne C. et Gaffuri J. 2008. *Combining three multi-agent based generalisation models : AGENT, CartACom and GAEL*. In Proceedings of the 13th International Symposium on Spatial Data Handling, Montpellier, France, 2008, p.277-296.

Gleyze J.-F., 2008, *Using structural approach to understand transportation networks vulnerability*. Proceedings of European Geosciences Union 2008, Vienna, Austria

Olteanu-Raimond A.M., Mustière S. 2008. *Data matching - a matter of belief*. Proceedings of the International Symposium on Spatial Data Handling SDH'2008, 23-25 juin, Montpellier, 2008, p. 501-519.

Taillandier P., Duchêne C., Drogoul A. 2008. *Knowledge revision in systems based on an informed tree search strategy: application to cartographic generalisation*, International Conference on Soft Computing as Transdisciplinary Science and Technology (CSTST).

Touya G. 2008b. *Enrichissement de données par analyse spatiale pour la généralisation des réseaux*. Colloque International de Géomatique et d'Analyse Spatiale SAGEO'08, Montpellier, France.

2007

Beguec J., Lafaye J., Gross-Amblard D., Ruas A., 2007, *Une méthode de tatouage du bâti robuste à l'équarrissage*, SAGEO, Clermont Ferrand, 18-20 juin 2007.

Béguec J., Lafaye J., Gross-Amblard D. and Ruas A., 2007, *Invisible Graffiti on your Buildings: Blind & Squaring-proof Watermarking of Geographical Databases*, in 10th International Symposium on Spatial and Temporal Databases (SSTD07), Boston, MA July 16-18, LNCS 4605, pp. 312-329

Bucher B., Jolivet L., Buard E., Ruas A., 2007, *The need for Web legend services*, in J.M.Ware and G.E.Taylor (Eds), proceedings of the 7th International Symposium on Web and Wireless GIS (W2GIS), Springer Lecture Note in Computer Science, Cardiff, UK

Gaffuri J., 2007, *Field deformation in an agent-based generalisation model: the GAEL model*, GI-days 2007 - young researches forum, volume 30, pages 1-24, September 2007

Jolivet L., Buard E., Bucher B., Ruas A., 2007, *Amélioration de légende sur le Web*, SAGEO, Clermont Ferrand, 18-20 juin 2007.

Olteanu A-M., 2007, *A multi-criteria fusion approach for geographical data matching*, In proceedings from 5th ISSDQ, 13-15 Juin, Enschede, Pays-Bas

Olteanu A-M., 2007, *Appariement de données géographiques utilisant la Théorie de l'Evidence*, dans les actes de la conférence LFA (La logique floue et ses applications), Nîmes 22-23 novembre 2007.

2006

Domingues C., Bucher B., 2006, *Application d'aide à la conception de légende*, actes du colloque SAGEO, Strasbourg, septembre 2006

Olteanu A.M., Mustière S., Ruas A., Desveaux E. 2006. *L'apport de données spatiales pour une base de données ethnographique*. Dans les actes du Colloque International de Géomatique et d'Analyse Spatiale (SAGEO'2006), Strasbourg, Septembre 2006.

Poupeau B., Bonin O., 2006, *3D analysis with high-level primitives: a crystallographic approach*, proceedings de SDH'06, Vienna.

Taillandier P., 2007, *Acquisition automatique de connaissances de guidage d'un processus de généralisation de données géographiques*, RJCIA, Grenoble.

Taillandier P., 2007, *Révision à base d'agents des connaissances de guidage d'un processus de généralisation de données géographiques*, JFSMA, Carcassonne.

CONFÉRENCES B (*Conférence avec sélection sur articles courts / Atelier avec sélection / Poster des conférences A – Actes publiés*)

2008

Bucher B., Jolivet L., 2008, *Acquiring service oriented descriptions of GI processing software from experts*. in Proceedings of the 11th AGILE Conference, Girona, Espagne

Burghardt D., Schmid S., Duchêne C., Stöter J., Baella B., Regnauld N., Touya G., 2008, *Methodologies for the evaluation of generalised data derived with commercial available generalisation systems*. 11th ICA Workshop on progress in automated map generalisation, Montpellier, 2008, disponible sur le site web de la commission en généralisation de l'ACI : http://aci.ign.fr/montpellier2008/papers/28_Burghardt_et_al.pdf.

Christophe S., (2008), *Making legends by means of painters' palettes*. In proceedings of the "Art and Cartography" Symposium - January 31st to February 2nd, Vienna, Austria, 2008 (et Lecture Notes in Geoinformation and Cartography series. Eds Springer (publication en 2009))

Christophe, S. 2008, *Creative Cartography based on dialogue*, in proceedings of AutoCarto 2008 conference, Shepherdstown (WV), USA, 8-11 September 2008.

Dogru O., Duchêne C., Mustière S., Ulugtekin N. 2008. *User Centric Mapping for Car Navigation Systems*. Proceedings of the 5th International Symposium on LBS and TeleCartography, Salzburg, November 2008.

Dogru O., Duchêne C., Mustière S., Ulugtekin N. *User Centric Mapping for Car Navigation Systems*. 11th ICA Workshop on progress in automated map generalisation, Montpellier, 2008, disponible sur le site web de la commission en généralisation de l'ACI : http://aci.ign.fr/montpellier2008/papers/02_Dogru_et_al.pdf.

Dominguès, C. (2008). *Description de cartes géographiques*. 8èmes journées francophones Extraction et gestion des connaissances (EGC2008) – Atelier sur la modélisation utilisateur et la personnalisation d'interfaces web. Sophia-Antipolis.

Gaffuri J., Duchêne C., Ruas A. 2008. *Object-field relationships modelling in an agent-based generalisation model*. 11th ICA Workshop on progress in automated map generalisation, Montpellier, 2008, site web de la commission en généralisation de l'ACI http://aci.ign.fr/montpellier2008/papers/16_Gaffuri_et_al.pdf.

Jolivet, L., *On-demand map design based on user-oriented specifications*. in proceedings of AutoCarto 2008 conference, Shepherdstown (WV), USA, 8-11 September 2008.

Nakos B., Mustière S., Gaffuri J., 2008. *A transition from simplification to generalisation of natural occurring lines*, 11th ICA Workshop on progress in automated map generalisation, Montpellier, 2008, disponible sur le site web de la commission en généralisation de l'ACI http://aci.ign.fr/montpellier2008/papers/25_Nakos_et_al.pdf.

Paget A., Perret J., Gleyze J.-F., 2008, *L'analyse des graphes hydrographiques pour la caractérisation du relief*. Actes de la conférence SAGEO, 4 p.

Stöter J., Anders K.-H., Baella B., Burghardt D., Davila F., Duchêne C., Pla M., Regnauld N., Rosenstand P., Schmid S., Touya G., Uitermark H., 2008. *A study on the state-of-the-art in automated map generalisation implemented in commercial out-of-the-box software*. 11th ICA Workshop on progress in automated map generalisation, Montpellier, 2008, disponible sur le site web de la commission en généralisation de l'ACI : http://aci.ign.fr/montpellier2008/papers/13_Stoter_et_al.pdf.

Taillandier P. 2008, *Diagnosis in systems based on an informed tree search strategy: application to cartographic generalisation*, CSTST Student Workshop.

Touya G. 2008a. *First Thoughts for the Orchestration of Generalisation Methods on Heterogeneous Landscapes*. 11th ICA Workshop on progress in automated map generalisation, Montpellier, 2008, disponible sur le site web de la commission en généralisation de l'ACI : http://aci.ign.fr/montpellier2008/papers/01_Touya.pdf.

2007

Balley S., 2007. *Why and how to describe spatial datasets structures?* 23th ICA conference, 4-10 August 2007, Moscow

Buard E., 2007, *Cartography of Human Wildlife Conflicts in Kenya*, 23rd ICA conference, 4-10 August 07, Moscow.

Buard E., Ruas A., 2007, *Evaluation of colour contrasts by means of expert knowledge for on-demand mapping*, 23rd ICA conference, 4-10 August 07, Moscow.

- Bucher B., 2007, *A cartographic Web client to explore remote GI methods*, in proceedings of the 23rd ICA conference, Moscow, Russia
- Bucher B., Balley S., 2007, *A generic preprocessing service for more usable data processing services*. in proceedings of the 10th AGILE Conference, Aalborg, Danemark
- Chesneau E., 2007, *Etude des contrastes de couleur pour améliorer la lisibilité des cartes*. In 8ème rencontres de ThéoQuant, Besançon, 10 au 12 janvier 2007, 19 pages.
- Chesneau E., 2007, *Improvement of Colour Contrasts in Maps : Application to Risk Maps*. In Actes de la 10ème Conférence Internationale en Science de l'Information Géographique AGILE, Aalborg, Danemark, 8 au 11 mai 2007, 14 pages.
- Christophe S., 2007, *Legend design by a novice user on the Web-* in proceedings of the 5th Geographic Information Days (GIDAYS'07) - 10-12 September 2007, Münster, Germany, 2007
- Christophe S., Bucher B., Ruas A., 2007, *A dialogue application for creative portrayal*, in proceedings of the ICA conference, Moscou, Russia
- Dominguès C., 2007, *Du besoin de l'utilisateur à la spécification de sa carte* (poster), In Actes des 18èmes journées francophones d'ingénierie des connaissances (IC'2007), p.341-342 : Cépaduès-Éditions.
- Gaffuri J., 2007, *Outflow preservation of the hydrographic network on the relief in map generalisation*, ICA conference, August 07, Moscow, Russia
- Gleyze J.-F., 2007, *Reducing functional network vulnerability - Structural indicators to anticipate the structural reorganisation of a damaged transportation network*, Actes de la 15e édition de la Conférence Européenne de Géographie Théorique et Quantitative, ECTQG, Montreux, Suisse, 4 p
- Olteanu A.-M., 2007, *Matching geographical data using the Theory of Evidence*, In proceedings from ICC, August 07, Moscow, Russia
- Pierkot C., Mustière S., 2007. *Gestion des mises à jour concurrentes dans des jeux de données géographiques répartis*. Actes de SAGEO'2007, papier court (4p), sur CDROM.
- Poupeau B., Ruas A., 2007, *A crystallographic approach to simplify 3D building*, in proceeding of the 23rd ICA conference, Moscow, Russia
- Taillandier P., 2007, *Automatic Knowledge Revision of a Generalisation System*, Workshop in generalisation and multiple representation, Moscow.
- Touya G., 2007, *A Road Network Selection Process Based on Data Enrichment and Structure Detection*. In proceedings of 10th ICA Workshop on Generalisation and Multiple Representation. Moscow (Russia).
- Touya G., 2007, *River Network Selection based on Structure and Pattern Recognition*. In proceedings of ICC2007, the International Cartography Association Conference, Moscow (Russia).

2006

- Abadie N., Gesbert N., Mustière S., 2006, *Création d'une ontologie à partir des spécifications textuelles pour l'intégration des bases de données géographiques* (papier court). Dans les actes de la conférences Ingénierie des Connaissances (IC'2006), Nantes, Juin 2006.
- Abd El Kader Y., Bucher B., 2006, *Cataloguing GI Functions provided by Non Web Services Software resources Within IGN*, in proceedings of the AGILE conference, Budapest, April 2006
- Balley S., 2006, *A web interface to explore and restructure a dataset*, in proceedings of the GISRUK conference, Nottingham, April 2006. 3rd prize for the best young researcher paper.
- Balley S., Bucher B., Libourel T., 2006, *A service to customize the structure of a geographical dataset*, Semantic Based GIS workshop (SEBGIS), Montpellier, November 2006
- Bonin O. et Poupeau B., 2006, *Raisonnement spatial sur la position et l'orientation de solides dans l'espace*, proceedings de l'atelier RTE'06, Nantes.
- Bonin O., 2006, *Sensitivity analysis and uncertainty analysis for vector geographical applications*, proceedings de Accuracy'06, Lisboa
- Bonin O., Deffontaines B., Poupeau B., Gueguen Y. and Fruneau B., *Combining ground displacement information with a 3D geographical information system to perform post-mining diagnosis*, accepted to the ISPRS Commission VIII conference, Haïfa, Israel, 2006.
- Chesneau E., 2006, *Un système pour améliorer de façon automatisée la représentation cartographique du risque*, revue MOSELLA, Actes du colloque Spatialisation et cartographie en Hydrologie, Université de Metz, 8-10 septembre 2004, 10 pages
- Domingues C., Bucher B., 2006, *Legend design based on map samples*, in proceedings of the 4th GIS Science conference, Extended abstracts, Raubal, Miller, Frank and Goodchild (Eds), IFGI Prints, Münster, September 2006
- Gaffuri J., 2006. *Deformation using agents for map generalization - Application to the preservation of relationships between fields and objects*, AutoCarto'06, Vancouver, United-states, 2006

- Gaffuri J., 2006. *How to merge optimization and agent-based techniques in a single generalization model?*, workshop in generalisation and multiple representation, Vancouver, 2006.
- Gaffuri J., 2006. *Système multi-agent pour la déformation en généralisation cartographique*, Journées Francophones sur les Systèmes Multi-Agents JFSMA 2004, Annecy, France, 2006.
- Olteanu A-M., Mustière S., Ruas A., *Matching Imperfect Data*, Proceedings from 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Lisbon 7-9 July 2006, pp 694-704.
- Pierkot C., Mustière S., Ruas A., Hameurlain A., 2006, *Using metadata to help the integration of several multi-sources set of updates*. 9th International Conference for Global Spatial Data Infrastructure (GSDI'06), Santiago du Chili, nov 2006.
- Poupeau B., Bonin O., 2006, *Cristage: a 3D GIS with a logical crystallographic layer to enable complex analysis*, in proceedings of 3DGeoInfo, Kuala Lumpur, Malaysia, July 2006
- Ruas A., Dyevre A., Duchêne C., Taillandier P., 2006, *Methods for improving and updating the knowledge of a generalization system*, Autocarto 2006, Portland USA.
- Touya G. 2006. *A Method for Generalization of River Networks Based on "Strokes" and Database Enrichment*. Extended Abstracts Proceedings of 4th International Conference GIScience 2006, Münster, Germany. pp 191-194.
- Touya G., Grosso E., Mustière S., Duchêne C., 2006. *Generalization or Data Matching for Integration? Insights from Experiments on Databases at Different Levels of Details*. In Proceedings of the 4th International Conference on Geographic Information Science (GIScience'2006) – Extended Abstracts. Raubal, Miller, Frank and Goodchild (eds), IFGI Prints, vol.28, pp.373-376

AUTRES COMMUNICATIONS

2008

- Bucher, B., Jolivet, L., *Des services cartographiques additionnels pour des cartes sur mesure efficaces*. Géo-événement, Paris, 2008
- Christophe, S. (2008) *Concevoir des légendes topographiques à l'aide de palettes de peintres*, Assemblée Générale du Comité Français de Cartographie (CFC), Jeudi 22 Mai 2008, Musée de la Marine, Paris. "Le monde des Cartes", CFC, N°196, Juin 2008.
- Gleyze J.-F., 2008, *Évaluation structurelle des réseaux de transport en géographie quantitative – Exemples d'application*. Journée de travail « Réflexion sur la modélisation des réseaux routiers en archéologie » du projet Dynarif – Dynamique et résilience des réseaux routiers et parcellaires en région Ile-de-France, Nanterre, 12 février 2008
- Gleyze J.-F., 2008, *L'approche structurelle des réseaux de transport – Application à l'étude de la vulnérabilité – Simplification des réseaux par l'étude des dépendances relationnelles*. Journées de travail du projet SPANGEO – Multi-level modelling of Spatial Networks in Geography, Cargèse, France, 30 janvier – 2 février 2008
- Ruas A., 2008, *Continuité et modélisation des données. Les travaux COGIT et le projet GeOpenSim*. Atelier PUCA Cohérence, continuité de l'information en regard de la disparité et de la continuité des territoires Avril 2008.

2007

- Abadie N., Olteanu A-M. et Mustière S., 2007, *Comparaison de la nature d'objets géographiques*, dans les actes de la conférence Ingénierie des connaissances, la journée OGHS « Ontologies et Gestion de l'Hétérogénéité Sémantique », Grenoble, 3 juillet 2007
- Balley S., 2007, *Améliorer l'interopérabilité entre données et traitements - Approches du laboratoire COGIT*. Présentation à la Journée Française de l'Interopérabilité (<http://geo-interop.brgm.fr>), réunion technique OGC, Enghien-les-Bains, juillet 2007.
- Balley S., 2007, *Pourquoi et comment décrire la structure d'un jeu de données géographiques*, numéro spécial de la revue "Le monde des cartes" du comité français de cartographie, sur les publications à la conférence ICC'07, n°194, décembre 2007
- Buard E., 2007, *Evaluation des contrastes colorés en cartographie par expertise*, numéro spécial de la revue "Le monde des cartes" du comité français de cartographie, sur les publications à la conférence ICC'07, n°194, décembre 2007
- Bucher B., 2007, *La Carte à la carte sur le Web*, dans Le Monde des Cartes, revue du comité français de cartographie, No194
- Bucher B., 2007, *Le projet européen SPIRIT : un moteur de recherche géographique d'information sur le Web*, Bulletin de la recherche de l'IGN 07
- Bucher B., 2007, *Un client Web pour découvrir, explorer et tester des traitements sur des données géographiques*, numéro spécial de la revue "Le monde des cartes" du comité français de cartographie, sur les publications à la conférence ICC'07, n°194, décembre 2007
- Christophe S., Bucher B., Ruas A., 2007, *Une application de dialogue pour des légendes créatives*, numéro spécial de la revue "Le monde des cartes" du comité français de cartographie, sur les publications à la conférence ICC'07, n°194, décembre 2007
- Gaffuri J., 2007, *Déformation pour la généralisation à base d'agents: le modèle GAEL*, Bulletin d'information scientifique et technique, Institut géographique national, 9 octobre 2007

- Gaffuri J., 2007, *Préservation de l'écoulement du réseau hydrographique sur le relief en généralisation cartographique*, numéro spécial de la revue "Le monde des cartes" du comité français de cartographie, sur les publications à la conférence ICC'07, n°194, décembre 2007
- Mustière S., 2007, *Recherches en intégration de bases données géographiques au laboratoire COGIT de l'IGN*, Présentation au GéoÉvénement
- Mustière S., Abadie N., Laurens F., 2007, *Appariement de schémas de BD géographiques à l'aide d'ontologies déduites des spécifications*. Actes de l'atelier DECOR (passage à l'échelle des techniques de découverte de correspondances), Namur - Belgique, janvier 2007 pp.22-27.
- Olteanu A.-M., 2007, *Appariement de données géographiques utilisant la théorie des croyances*, numéro spécial de la revue "Le monde des cartes" du comité français de cartographie, sur les publications à la conférence ICC'07, n°194, décembre 2007
- Poupeau B., 2007, *Utilisation de graphes en 3D pour modéliser les relations d'adjacence entre objets géographiques et leurs évolutions dans le temps*, journée de présentations et d'échanges autour des applications des graphes en géographie et en analyse spatiale
- Taillandier P., 2007, *Révision des connaissances d'un processus de généralisation de données géographiques*, numéro spécial de la revue "Le monde des cartes" du comité français de cartographie, sur les publications à la conférence ICC'07, n°194, décembre 2007
- Touya G. 2007. *Un processus de sélection du réseau hydrographique basé sur la détection des structures*, numéro spécial de la revue "Le monde des cartes" du comité français de cartographie, sur les publications à la conférence ICC'07, n°194, décembre 2007

2006

- Bordin P., *De la carte à l'ordinateur pour observer le territoire*, numéro spécial de la Recherche "Ville et mobilité durable", n° 398, juin 2006
- Chesneau E., 2006, *Pour une amélioration automatique des contrastes colorés en cartographie : application aux cartes de risques*, dans le bulletin du Comité français de Cartographie, décembre 2006, 16 pages
- COGIT, 2006, CAC105, Cartographie des activités de recherche du laboratoire Cogit 2001-2005.
- Gesbert N., 2006, *Formalisation des spécifications des bases de données géographiques*, Bulletin d'Information Scientifique et Technique de l'IGN, n°75, Recherches 2003, pp.81-86
- Gesbert, Nils. 2006. *Formalisation des spécifications des bases de données géographiques*. Bulletin d'information scientifique et technique de l'IGN, no 75, p. 81-86.
- Gleyze J-F., 2006, *Extraction de l'information implicite contenue dans les bases de données géographiques - Application à l'étude de la vulnérabilité des réseaux de transport*, Séminaire du laboratoire LABRI, Université de Bordeaux I, 24 janvier 2006
- Gleyze J-F., 2006, *Extraction de l'information implicite contenue dans les bases de données géographiques - Application à l'étude de la vulnérabilité entre les chercheurs en socio-économie des transports et le réseau du Tourisme et de la Mer*, grande Arche de la Défense, 30 juin 2006, 14p
- Gleyze J-F., 2006, *L'apport de l'information géographique dans l'analyse des risques - Application à l'étude de la vulnérabilité des réseaux de transport*, Séminaire du GDR Sigma 2 consacré aux risques, Institut de Géographie de Paris, 14 juin 2006
- Gleyze J-F., 2006, *Les outils de connaissance des réseaux - Approche par la vulnérabilité pour les réseaux de transport*, Journée scientifique et technique ENGEES - CEMAGREF "La gestion patrimoniale des équipements d'eau potable et d'assainissement - décideurs, gestionnaires et chercheurs travaillant en réseau, ENGEES, Strasbourg, 17 janvier 2006
- Mustière S. 2006. *Results on experiments on automated matching of networks*. Proceedings of the ISPRS Workshop on Multiple Representation and Interoperability of Spatial Data, Hanover, pp.92-100
- Mustière S., 2006, *Consistent integration of geographic databases: examples and tools*, INSPIRE workshop on data consistency, Ispra, Italy, November 2006
- Ruas A., 2006, *Recherche en automatisation des processus cartographiques*, Journées CFC
- Ruas A., 2006, *Traitements de l'information géographique et recherche en SIG*, Les Imagies Montpellier, 27 octobre 2006
- Ruas A., 2006, *Usages des bases de données géographiques- Exemple de l'IGN*, France, Rencontres Jacques Cartier, 4-6 décembre 2006, Lyon
- Sheeren D. 2006. *Étude de la cohérence inter-représentations : vers une meilleure intégration des bases de données spatiales*. Bulletin d'information n.75 de l'IGN. pp.71-80.

RAPPORTS TECHNIQUES, RAPPORTS DE STAGE

Rapports techniques (sélection) :

- Brasebin M., Ruas A., Velten A. et G. Touya – Tâche WP4 *Implémentation d'un algorithme de généralisation automatique de bâtiment 3D*. TerraNumérica.

- Bucher B., WP 3.1.3.2 – D2, *Spécifications de services d'interface de contenu TerraNumerica*, 2008
- Duchêne C., Gaffuri J. *Description des méthodes d'analyse de la ville implémentées au COGIT* Rapport technique 2.2 du projet GeOpenSim, 35 pages.
- Mustière S., Duchêne C., Touya G. 2006. *Généralisation et intégration du fond vert*. Rapport technique IGN, SR/2006.0369
- Paget A., 2008, *Apport de la caractérisation du relief pour des applications risque*. Rapport technique IGN, 200 Pages. IGN-DT/2008-191
- Perret J., 2008, *Création de bases de données temporelles pour le projet GeOpenSim*, Rapport Technique du projet GeOpenSim
- Perret J., 2008, *Modélisation des agents dans GeOpenSim*, Rapport Technique 1.1 du projet GeOpenSim.
- Perret J., 2008, *Modélisation fonctionnelle du système GeOpenSim*, Rapport Technique 1.2 du projet GeOpenSim.
- Touya G., Grosso E., Mustière S., Duchêne C. 2006. *Automatisation de l'appariement de la BDCarto avec la BDTopo Pays et de la généralisation de la BDCarto à partir de la BDTopo Pays, en vue de l'intégration de ces bases - « La BDCARTO fait peau neuve »*. Rapport technique interne IGN, ref SR/2006.0045.

Rapports de stage :

2008

- Baldit Priscille, 2008, *Exploitation (avec des outils statistiques et de TALN) d'un corpus d'interviews sur le thème de la randonnée* M2 Traductique et Gestion de l'Information à l'Institut national des Langues et Civilisations Orientales (INaLCO) – Paris
- Gonindard Nicolas, 2008, *Tests d'évaluation des contrastes colorés*. Cycle d'ingénieur ENSG 3^{ème} Année. Master Carthageo-Pro. 63 pages
- Horel Thomas, 2008. *Interface de consultation des spécifications de bases de données géographiques*. Stage de 2^{ème} année d'école d'ingénieurs de l'ESTIA (Ecole Supérieure des Technologies Industrielles Avancées).
- Minard Anne-Lyse, 2008. *Recherche et analyse de ressources terminologiques liées à la topographie*. Rapport de stage de Master 1 « Traitement automatique des langues », Université Lille 3, soutenu en septembre 2008.
- Renard, Jérémie (2008). *Caractérisation et analyse des légendes des cartes topographiques européennes*. Rapport de stage de mastère Carthageo de l'Université Paris 1, Paris 7 et de l'ENSG.
- Taibi Salima, 2008, *Système multi-Agents sur un Cercle Chromatique* Master 2 informatique, Université Paris 5 (René Descartes) spécialité Systèmes Multi-Agents Stage soutenu le 29 juin 2008

2007

- Ménégaux Guillaume, 2007, *Consultation et classification d'une base d'échantillons cartographiques*. Rapport de Stage de Mastère 2 en Sciences de l'Information Géographique, Ecole Nationale des Sciences Géographiques
- Picard V. 2007. *Instanciation automatique des liens entre ontologies et schémas de bases de données géographiques à partir des spécifications en langage naturel*. Master2 DEFI (Documents Electroniques et Flux d'Information), université Paris X - Nanterre.
- Velten Aurélien, 2007, *Implémentation d'un algorithme de généralisation automatique de bâtiments 3D*. Rapport de stage M2 Sciences de l'Information Géographique, Université Marne-la-Vallée. 44 pp

2006

- Aramini Brice, 2006, *Conception d'une interface de manipulation de composants fonctionnels portant sur des données géographiques*. Rapport de Stage de Mastère Informatique appliquée aux Systèmes d'Information Géographique, Université Pierre et Marie Curie
- Béguec Jean, 2006, *Diffusion et tatouage de données géographiques*. Rapport de stage de Master 2 SIG de l'Université Marne-la-Vallée et de l'ENSG, en partenariat avec le CNAM. 29 pages
- Culson Nicolas, 2006, *Détection et généralisation des échangeurs routiers dans une base de données cartographiques vectorielles*. Rapport de stage M2 Informatique Appliquée aux Systèmes d'Information Géographique, Université Pierre et Marie Curie, Paris 6. 85 pp.
- Laurens F. 2006. *Création d'une ontologie à partir de textes en langage naturel*. Stage de Master 1 Linguistique-Informatique, Université Paris 7.
- Philippe Pierre, 2006, *Définition d'un jeu d'échantillons cartographiques qualifiés*. Rapport de Stage de Mastère Carthageo Pro de l'Ecole Nationale des Sciences de l'information Géographique
- Teng J. 2006. *Conception et réalisation d'outils ergonomiques pour la navigation au sein de bases de données géographiques à représentation multiple*. Master 2 Informatique Appliquée aux Systèmes d'Information Géographiques (M2 IASIG), université Paris 6 et ENSG
- Vilars Xavier, 2006, *Propagation d'opérations de généralisation de l'hydrographie sur le relief*, Master ASIG, Université Marne-La-Vallée