Formalisation d'un environnement d'aide à l'analyse géovisuelle

Gabriel Vatin

▶ **To cite this version:**

Gabriel Vatin. Formalisation d'un environnement d'aide à l'analyse géovisuelle : Application à la sécurité et sûreté de la maritimisation de l'énergie. Environnements Informatiques pour l'Apprentissage Humain. Ecole Nationale Supérieure des Mines de Paris; CRC - Centre de recherche sur les Risques et les Crises, 2014. Français. NNT : . tel-01253688

HAL Id: tel-01253688

https://theses.hal.science/tel-01253688

Submitted on 11 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

MINES ParisTech

École doctorale n° 432 : Sciences des Métiers de l'Ingénieur

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

l'École nationale supérieure des mines de Paris

Spécialité " Sciences et Génie des Activités à Risques "

présentée et soutenue publiquement par

Gabriel VATIN

le 21 Novembre 2014

Formalisation d'un environnement d'aide à l'analyse géovisuelle

Application à la sécurité et sûreté de la maritimisation de l'énergie

Directeur de thèse : **Aldo NAPOLI**

Jury

Mme Sara FABRIKANT, Professeur, Dept Geography, University of Zurich	Rapporteur
M. Rodolphe DEVILLERS, Professeur, Dept Geography, Memorial University of Newfoundland	Rapporteur
M. Christophe CLARAMUNT, Professeur, IRENAV, École Navale	Examinateur
M. Thierry JOLIVEAU, Professeur, UFR SHS, Université Jean Monnet Saint-Etienne	Examinateur
M. Franck GUARNIERI, Directeur de recherche, CRC, MINES ParisTech	Examinateur
M. Aldo NAPOLI, Chargé de recherche, CRC, MINES ParisTech	Examinateur

T H È S E

MINES ParisTech
Centre de recherche sur les Risques et les Crises
Rue Claude Daunesse, CS 10207 - 06904 Sophia Antipolis Cedex - France

« Le savoir ne naissait jamais simplement dans l'esprit humain, il renaissait toujours.
La transmission de la sagesse d'un âge à l'autre, ce cycle des renaissances,
telle était la vraie sagesse. Le reste n'était que barbarie. »

Salman Rusdhie, **L'enchanteresse de Florence**

« So tonight you better stop and rebuild all your ruins
Because peace and trust can win the day
Despite of all your losing. »

Led Zeppelin, **Immigrant Song**

« We would shout and swim about
The coral that lies beneath the waves
Oh what joy for every girl and boy
Knowing they're happy and they're safe »

The Beatles, **Octopus's Garden**

Remerciements

Je tiens tout d'abord à remercier l'ensemble des membres du jury d'avoir accepté de prendre part à l'évaluation de cette thèse. La diversité de leurs horizons géographiques et scientifiques vient soutenir l'importance de notre domaine de recherche, et permet ainsi d'améliorer notre réflexion par des points de vue variés. Pour cela, j'adresse un profond merci à Rodolphe Devillers du Canada, Sara Fabrikant de Suisse pour avoir été rapporteurs. Je remercie également Christophe Claramunt, Thierry Joliveau et Franck Guarnieri d'avoir été les examinateurs de cette thèse, en dehors du CRC.

Evidemment, après avoir travaillé pendant trois années à ses côtés et sous sa direction, j'adresse une très grande reconnaissance et d'immenses remerciements à Aldo Napoli. Guider un tout jeune gradué d'école d'ingénieur à travers les chemins complexes et sinueux de la recherche doit être une grande épreuve, que l'on relève et améliore année après année et qui demande de la patience. Mais il a toujours été là pour me soutenir, me conseiller, me motiver et me pousser à aller plus loin. Mais avant tout, Aldo a été à l'écoute de toutes les propositions, aussi bien farfelues qu'elles puissent être. Certains doctorants peuvent se plaindre de ne pas avoir assez de liberté, et d'autres peuvent critiquer une absence de leur directeur : l'encadrement qu'il m'a offert au CRC est irréprochable et m'a toujours permis d'évoluer dans une communauté motivée, motivante et passionnante.

Au cours de ma période de jeune chercheur, j'aurais pu voir passer beaucoup de personnes dans mon entourage direct, et qui auront eu une grande influence sur mon travail, et ma vie professionnelle. Dernier arrivé dans une belle équipe de chercheurs dans le domaine maritime, j'aurais fini cette thèse en les voyant partir un à un vers de nouveaux horizons. Pour leur soutien professionnel, technique et surtout amical pendant cette vie de « laborantin », je remercie Arnaud, Xavier, Bilal et Amal. Pour leur contribution à nos nombreux projets, mes remerciements vont aussi vers les stagiaires qui ont pu passer quelques mois dans notre bureau, à savoir Adnan, Adrian, Thomas, Vincent, Jean-Eude et Matthieu.

Mais une thèse, c'est une simple petite brique dans une construction bien plus grande, qu'est le centre de recherches. Les enseignements et conseils procurés au CRC, du début de la première année à la fin de la dernière année, auront tous été d'une très grande utilité pour cette thèse, et pour ma carrière future. Pour cela je remercie Franck, directeur du centre, pour son accueil au CRC. De plus, le personnel enseignant qui m'aura accompagné et soutenu au cours des trois années, Eric, Jean-Luc, Emmanuel, Wim et Valérie. Je n'oublie pas le personnel informatique et administratif qui aura tant fait pour soutenir et simplifier le travail des doctorants sur le site de Sophia : Samuel, Sandrine, Myriam, Sylvie, Brigitte et Isabelle. Beaucoup de leur temps aura été attribué à notre confort de travail, et l'expérience de thèse aurait été toute autre sans leur soutien. Merci !

Enfin, au cours des nombreux déplacements auxquels j'ai pris part, j'ai pu rencontrer des personnes extraordinaires, aussi bien sur le point de vue scientifique qu'humain. Pour cela, mes derniers remerciements iront à l'ensemble de ces personnes de la communauté de l'information géographique, notamment de l'ITC aux Pays-Bas et de l'IGN en France. Pour le financement de cette thèse et le soutien technique apporté dans le développement de nos outils, je remercie finalement Alain Orengo, PDG de la société ITE, ainsi que la région PACA.

Table des matières

INTRODUCTION 1

Contexte : maritimisation de l'énergie 1
L'espace maritime, espace de rigueur, de liberté et de risques 1
Les acteurs de la « Sécurité et Sûreté de la Maritimisation de l'Energie » 3
Les outils pour la navigation et la surveillance maritime 5
Analyse et détection des comportements à risques : les systèmes de surveillance de nouvelle génération 7

Problématique de recherche : aide à l'analyse géovisuelle des données de mouvement 10
La visualisation, une approche anthropocentrique pour la prise de décision 10
L'espace géographique, support de l'analyse des mouvements 11
Une approche formelle pour la géovisualisation 13
Vers un environnement d'aide à l'analyse géovisuelle 13

Hypothèses et objectifs 14

Démarche de recherche 16

Structure de la thèse 16

CHAPITRE 1 - L'ETRE HUMAIN FACE AUX COMPORTEMENTS A RISQUES 19

1.1. De la gestion des risques à l'aide à l'analyse 19
1.1.1. Les processus de gestion des risques 19
1.1.2. De l'aide à l'analyse à la prise de décision 31

1.2. L'utilisateur et l'usage des nouvelles technologies 39
1.2.1. Les limites de la cognition humaine face à la visualisation 40
1.2.2. L'acceptation de nouvelles technologies 42

1.3. Le rôle de l'information géographique pour la gestion des risques 51
1.3.1. Le processus de raisonnement par la carte 51
1.3.2. L'analyse des données de mouvement par la visualisation 54

Conclusion du Chapitre 1 59

CHAPITRE 2 - ANALYSE GEOVISUELLE ET FORMALISATION DE LA VISUALISATION 63

2.1. De la représentation de données à l'analyse géovisuelle 63
2.1.1. Perception de l'information 63
2.1.2. La cartographie, moyen représenter l'information connue 69
2.1.3. La visualisation d'information : explorer pour découvrir l'inconnu 77

2.2. Modélisation formelle des méthodes de visualisation d'information 86
2.2.1. Modélisation des données 86
2.2.2. Modélisation des méthodes de visualisation 89

2.2.3. Utilisation de formalisation pour une visualisation automatisée 96

2.3. Formaliser des connaissances : recours aux ontologies 98
2.3.1. Définition générale des ontologies 98
2.3.2. Constitution d'une ontologie 99
2.3.3. Raisonnement : utilisation de règles et inférence 104

Conclusion du Chapitre 2 108

CHAPITRE 3 - CONCEPTION D'UN ENVIRONNEMENT D'AIDE A L'ANALYSE GEOVISUELLE POUR LES RISQUES MARITIMES 111

3.1. Méthodologie proposée 111
3.1.1. Un environnement d'aide à l'analyse géovisuelle 112
3.1.2. Un système à base de connaissances pour l'analyse géovisuelle appliquée au domaine maritime 114
3.1.3. Méthodologie de développement de l'ontologie 117

3.2. Modélisation des concepts d'intérêt 120
3.2.1. Domaines couverts par l'ontologie 120
3.2.2. Modélisation du contexte d'utilisation : le domaine maritime et la situation d'utilisation 121
3.2.3. Modélisation des environnements d'analyse géovisuelle 136
3.2.4. Lien entre l'utilisateur et la visualisation 145
3.2.5. Ontologie globale et instances développées 145

3.3. Modélisation des règles d'inférence : recherche de solutions pour l'analyse géovisuelle 149
3.3.1. Plan d'exécution des règles 152
3.3.2. Ensemble de règles développées : processus de raisonnement 153

Conclusion du Chapitre 3 163

CHAPITRE 4 - EXEMPLE D'UTILISATION DE L'ENVIRONNEMENT D'AIDE A L'ANALYSE GEOVISUELLE 165

4.1. Plateforme web de géovisualisation maritime : FishEye 165
4.1.1. Rôle de la plateforme 165
4.1.2. Exemples de visualisations 167
4.1.3. Interface d'utilisation de l'environnement d'aide à l'analyse géovisuelle 170

4.2. Expérimentation du module 178
4.2.1. Amélioration de la visualisation du trafic maritime 178
4.2.2. Un cas d'étude : collision de navires 179
4.2.3. Exploration et analyse des mouvements de navires 184

4.3. Performance du raisonneur ontologique 197

4.4. Retour sur la gestion des risques 199

4.5. Perspective d'amélioration du module 200
4.5.1. Ordre de proposition des visualisations 200
4.5.2. Auto-évaluation du profil utilisateur 201

Conclusion du Chapitre 4 ..... **202**

CONCLUSION ET PERSPECTIVES **205**

Bilan de la thèse **205**
Définition de l'aide à l'analyse géovisuelle pour les utilisateurs 206
Conception d'un modèle pour la formalisation de la visualisation 206
Intégration dans un environnement d'analyse du trafic maritime 206
Ouverture à la résilience 206

Perspectives **207**
Perspectives de modélisation 207
Perspectives d'application 208

BIBLIOGRAPHIE **213**

Liste des acronymes

ACI	Association Cartographique Internationale
AEM	Action de l'Etat en Mer
AIS	*Automatic Identification System*
AMP	Aires marines protégées
ANR	Agence Nationale de la Recherche
ARPA	*Advanced Research Projects Agency*
BEAmer	Bureau enquêtes des accidents de mer
CeCLAD-M	Centre de Coordination de la Lutte Anti-drogue en Méditerranée
COCOM	*Contextual Control Model*
COP	*Common operational picture*, vue opérationnelle commune
CPA	*Closest Point of Approach*, point de plus proche approche
CROSS	Centre régional opérationnel de surveillance et de sauvetage
DGPS	*Differential Global Positioning System*
DSS	*Decision Support System*
EAAG	Environnement d'aide à l'analyse géovisuelle
ECDIS	*Electronic Chart Display and Information System*
ECOM	*Extended Control Model*
ECS	*Export Control System*
ENC	*Electronic Navigational Chart*
FAO	*Food and Agriculture Organization*
FCA	*Formal Concept Analysis*
FM-CW	*Frequency-Modulated Continuous-Wave*, radar à onde continue modulée en fréquence
GEBCO	*General Bathymetric Chart of the Oceans*
GeoVA	*GeoVisual Analytics*, analyse géovisuelle
GeoVis	Géovisualisation
GPS	*Global Positioning System*
HAJL	*Human-Automated Judge Learning*
HSF	*Hybrid Spatio-temporal Filtering*
IG	Information géographique
IGAM	Inspection Générale des Affaires Maritimes
IHM	Interface homme-machine
JCS	*Joint Cognitive System*
MAIB	*Marine Accident Investigation Branch*
MAUP	*Modifiable Areal Unit Problem*
MISR	*Maritime Intelligence, Surveillance & Reconnaissance*
MMSI	*Maritime Mobile Service Identity*

MSRAM	*Maritime Security Risk Analysis Model*
NMEA	*National Marine Electronics Association*
OHI / IHO	Organisation Hydrographique Internationale, *International Hydrographic Organization*
OMI / IMO	Organisation Maritime Internationale, *International Maritime Organization*
PEOU	*Perceived Ease-Of-Use*
PSE	*Problem Solving Environment*
PU	*Perceived Usefulness*
RADAR	*Radio Detection and Ranging*
SA	*Situation Awareness*
SAD	Système d'aide à la décision
SAR	*Search and Rescue*
SARDS	Système d'aide à la décision à référence spatiale
SHOM	Service hydrographique et océanographique de la marine
SI	Système d'information
SIG	Système d'information géographique
SOM	*Self-Organizing Map*, carte auto-adaptative
SSM	Système de surveillance maritime
SSME	Sécurité et sûreté de la maritimisation de l'énergie
TAM	*Technology Acceptance Model*
TI	Technologies de l'information
UNCTAD	*United Nations Conference on Trade and Development*
VMS	*Vessel Monitoring System*
VTS	*Vessel Traffic Services*
W3C	*World Wide Web Consortium*
ZEE	Zone économique exclusive
ZPE	Zone de protection écologique
ZPP	Zone de pêche protégée

Liste des figures

Figure 0-1. Infographie de la localisation des attaques pirates en 2011 (*source : Travelinsurance.org*) 3
Figure 0-2. Organisation de l'*Action de l'Etat en mer*, en France 4
Figure 0-3. Répartition des CROSS en France métropolitaine (*source : marine-marchande.net*) 4
Figure 0-4. Ecran radar pour l'aide à la navigation maritime (*source : Wikimedia*) 5
Figure 0-5. Surcharge d'information sur un ECDIS (*source : nautinst.org*) 6
Figure 0-6. Amélioration des composantes des systèmes de surveillance (*source :* Vandecasteele 2012) 9
Figure 0-7. Evolution de la cartographie pendant le Moyen-Âge (*www.bimcc.org*) 12
Figure 1-1. Courbe de Farmer : gravité et probabilité des risques 20
Figure 1-2. Différentes stratégies de réduction des risques (d'après *United States Coast Guard*) 21
Figure 1-3. Boucle OODA, ou cycle de Boyd, pour la prise de décision dans le domaine militaire 23
Figure 1-4. Processus de maîtrise des risques (d'après Wybo 2012) 24
Figure 1-5. Modèlisation d'un système socio-technique et impacts des systèmes d'information sur ce système (d'après Kéfi et al. 2004) 27
Figure 1-6. Modèle COCOM pour le contrôle des risques (d'après Hollnagel 1993) 27
Figure 1-7. Modèle ECOM pour le contrôle des risques (d'après Hollnagel et Woods 2005) 28
Figure 1-8. Exemple de réseau bayésien à trois nœuds, ainsi que les probabilités correspondantes.. 34
Figure 1-9. Analyse des comportements de navires par fouille de données (d'après Maggi et al. 2013) 36
Figure 1-10. Comparaison des SIG et du SOLAP pour l'aide à la décision (d'après Rivest et al. 2001) 38
Figure 1-11. Opération de drill-down en SOLAP (*source : spatialolap.scg.ulaval.ca*) 38
Figure 1-12. ColorBrewer 2.0, proposition d'échelles de couleurs selon la classification des données41
Figure 1-13. Différents niveaux d'automatisation sur les fonctions lors de la prise de décision (d'après Parasuraman et al. 2000) 43
Figure 1-14. Interactivité / Automatisation : deux aspects complémentaires pour l'étude de données spatiales pour l'analyse des risques maritimes. Les exemples sont placés de manière approximative. 45
Figure 1-15. TAM : Modèle de l'acceptation de la technologie corrigé (d'après Davis 1993) 47
Figure 1-16. Modèle TAM 2, décrivant les facteurs d'influence sur l'utilité perçue 47
Figure 1-17. Modèle TAM 3, réutilisant les facteurs de TAM 2 (non placés ici) et les facteurs d'influence sur la facilité d'utilisation perçue 48
Figure 1-18. Modèle de la triade pour la définition d'un évènement (d'après Peuquet 1994)..... 51
Figure 1-19. Illustration du MAUP pour l'étude du transport routier pour quatre maillages différents (*source :* Viegas et al. 2009)..... 53
Figure 1-20. Processus d'analyse spatiale de données de mouvement (d'après Andrienko et al. 2011) 55
Figure 1-21. Processus d'analyse de trajectoires (d'après Andrienko et Andrienko 2013) 56
Figure 1-22. Définition de la MISR (d'après Davenport et Risley 2006)..... 57
Figure 1-23. Différents motifs récurrents en analyse de trajectoires (d'après Bogorny 2010) 58
Figure 2-1. Fonctionnement de la perception pré-attentive (*source :* Université François Rabelais) .. 64
Figure 2-2. Détection d'un objet distinct par la perception pré-attentive (*source :* Healey et al. 1996) 65
Figure 2-3. Vase de Rubin : contradiction sur l'interprétation des formes..... 66

Figure 2-4. Loi de proximité 66
Figure 2-5. Loi de clôture..... 66
Figure 2-6. Amas de choléra pendant l'épidémie de choléra de Broad Street par John Snow (1854) (*source : Wikimedia*)..... 70
Figure 2-7. Carte de la campagne Napoléonienne en Russie, par Minard (1869) (*source : Wikimedia*) 70
Figure 2-8. Cartes abstraites : (a) chorème de la région PACA (*source : Magali Brunet 2013*) et (b) cartogramme de surface de la population dans les états membres de l'UE (*source : Wikimedia*) 71
Figure 2-9. L'espace des utilisations de la carte : (Cartographie)^3 (d'après MacEachren, 1995)..... 73
Figure 2-10. Cartographie multimédia (d'après Antoni et al. 2004) 75
Figure 2-11. Schéma de la cognition humaine, dans le cas du raisonnement et de la prise de décision (*source :* Patterson et al. 2014) 78
Figure 2-12. Processus de visualisation (*source : InfoVis-Wiki*)..... 79
Figure 2-13. Processus de conception pour la visualisation d'information (*source :* Patterson et al. 2014)..... 79
Figure 2-14. Rôles de la géovisualisation (Kraak 2006) 81
Figure 2-15. Environnement d'exploration pour l'analyse géovisuelle (d'après Kraak 2012) 83
Figure 2-16. Technique du *brushing*, pour une sélection de longitudes moyennes (cadre de l'axe du milieu)..... 84
Figure 2-17. Utilisation de plusieurs liens entre les informations sélectionnées, sur plusieurs visualisations : flou, transparence, leader line (*source :* Robinson 2011)..... 85
Figure 2-18. Interface de l'outil Map4Decision pour l'analyse SOLAP (*source : Intelli3*) 85
Figure 2-19. Infographie pour présenter les différentes zones d'accidents maritimes et les facteurs de risques (*source : www.maritimepassivesafety.com*)..... 88
Figure 2-20. Classification des techniques de visualisation d'information (*source :* Keim 2002) 89
Figure 2-21. Reference Model Pattern pour la visualisation d'information (d'après Heer et Agrawala 2006)..... 90
Figure 2-22. Data State Reference Model pour la modélisation de la visualisation d'information (d'après Chi 2000) 92
Figure 2-23. Modélisation du processus de visualisation pour des histogrammes superposés (*source :* Daassi et al. 2005) 93
Figure 2-24. Intégration de la dimension temporelle en visualisation d'information (*source :* Daassi et al. 2005)..... 93
Figure 2-25. Symboles de Bertin pour la caractérisation de visualisations..... 94
Figure 2-26. Schématisation de Bertin pour la caractérisation d'un histogramme à axes parallèles (repris dans Hurter 2010)..... 94
Figure 2-27. Caractérisation de Wilkinson pour l'étude du taux de mortalité en fonction du taux de natalité, ainsi que leur densité (*source :* Wilkinson 1999, Hurter 2010)..... 95
Figure 2-28. Caractérisation de Baudel pour la visualisation des statistiques de criminalité dans les grandes villes américaines densité (*source :* Hurter 2010) 95
Figure 2-29. Architecture de la solution BDVR (*source :* Gotz et Wen 2009)..... 97
Figure 2-30. Base de connaissances à quatre concepts (bleu) et deux individus (orange)..... 100
Figure 2-31. Architecture du web sémantique, selon le W3C (d'après les propositions du W3C) 102
Figure 2-32. Approche hybride pour le développement d'une ontologie (d'après López-Pellicer et al. 2007)..... 102

Figure 2-33. Approche descendante, pour la conceptualisation des éléments de bas niveau, par la méthode Methontology (d'après Bendaoud et al. 2007) 103
Figure 2-34. Approche ascendante, pour la conceptualisation des éléments de haut niveau (*source :* López-Pellicer et al. 2007) 104
Figure 3-1. Mode d'utilisation d'un environnement d'aide à l'analyse géovisuelle 113
Figure 3-2. Place d'un EAAG au sein de la prise de décision, intégré dans le modèle COCOM 113
Figure 3-3. Modèle général pour l'étude des cas d'utilisation de l'analyse géovisuelle..... 116
Figure 3-4. Domaines ontologiques couverts pour l'analyse géovisuelle du trafic maritime 121
Figure 3-5. Hiérarchie de la typologie des données maritimes, dans l'éditeur Protégé 4.3 122
Figure 3-6. Caractérisation de la temporalité des données par les concepts et instances associées. 124
Figure 3-7. Catégorisation des anomalies maritimes, selon la taxonomie des faits (*source :* Roy et Danveport 2009)..... 125
Figure 3-8. Concept de risques et instances proposées dans l'ontologie 126
Figure 3-9. Typologie des tâches d'exploration et d'analyse 127
Figure 3-10. Cartographie des primitives de contrôle *orientées objectif* (*source :* Roth 2012) 128
Figure 3-11. Utilisation du concept de Closest Point of Approach pour la prédiction d'une collision entre deux navires (Cario et Gravière 2013) 129
Figure 3-12. Tâches d'identification générale pour l'étude du trafic maritime 130
Figure 3-13. Tâche d'analyse du contexte pour l'étude du trafic maritime..... 131
Figure 3-14. Tâches d'analyse spatio-temporelle pour l'étude du trafic maritime 131
Figure 3-15. Tâches d'identification du risque pour l'étude du trafic maritime 132
Figure 3-16. Le concept Goal, représenté par UML (a) et par une version sémantique (b) 132
Figure 3-17. Relations sémantiques entre le concept *User* et les caractéristiques de l'utilisateur 136
Figure 3-18. Modélisation UML pour la définition simple d'un environnement d'analyse géovisuelle 137
Figure 3-19. Classification générale des méthodes de visualisation d'information selon trois concepts généraux d'espaces géographiques, sémantiques et temporels 138
Figure 3-20. Relations sémantiques pour la description de la projection des données vers un espace visuel : concepts de données, types de données, formes visuelles et données visuelles 139
Figure 3-21. Cube spatio-temporel pour la visualisation d'une trajectoire dans l'espace et dans le temps : deux trajectoires de navires représentées dans le plan 2D et 3D..... 139
Figure 3-22. Instance du chemin spatio-temporel dans l'ontologie et liens sémantiques pour sa description..... 140
Figure 3-23. Concepts relatifs aux données et à leur projection visuelle, dans l'éditeur Protégé 4.3 : (a) les données visuelles, (b) les formes visuelles et (c) les données brutes 140
Figure 3-24. Formalisation de l'espace de visualisation par les relations entre concepts 142
Figure 3-25. Cartographie des primitives de contrôle *orientées opérateur* (*source :* Roth 2012) 144
Figure 3-26. Cartographie des primitives de contrôle *orientées opérande* (*source :* Roth 2012) 144
Figure 3-27. Concept Situation pour la description du cas d'utilisation et les propriétés correspondantes dans l'éditeur Protégé 4.3..... 145
Figure 3-28. Schéma global de l'ontologie développée, selon les concepts et leurs relations hiérarchiques et sémantiques 146
Figure 3-29. Formalisation des stratégies de visualisation pour deux scénarios : (a) analyse d'un comportement de pêche, (b) prédiction d'une collision 147

Figure 3-30. Formalisation de la méthode du cube spatio-temporel pour l'analyse d'une trajectoire dans l'espace et le temps 148
Figure 3-31. Liens sémantiques entre des instances de Situation, User et Goal 149
Figure 3-32. Intervention de l'ontologie et des règles pour la sélection de visualisations adaptées . 151
Figure 3-33. Copie des informations entre concepts, du plus bas niveau vers le plus haut 152
Figure 3-34. Suite logique pour la recherche de visualisations adaptées : parcours des différents concepts 153
Figure 3-35. Formalisme SWRL utilisé 153
Figure 3-36. Diagramme de communication maritime du Titanic, qui permet d'identifier les postes terrestres et les navires avec lesquels communiquer pendant la traversée 155
Figure 3-37. Visualisation non adaptée à la temporalité des données : étude de données temps réel par une visualisation type carte de Kohonen 158
Figure 3-38. Sélection de méthodes par les tâches proposées 159
Figure 3-39. Liste des espaces de visualisation adaptés à une situation données 162
Figure 4-1. Architecture d'une application N tiers pour la gestion d'informations géographiques ... 167
Figure 4-2. Barre d'outils de FishEye et sous-menus correspondants 167
Figure 4-3. Visualisation de la trajectoire d'un tanker en Mer Méditerranée 168
Figure 4-4. Visualisation de la trajectoire, vitesse, arrêts et informations sur un tanker 169
Figure 4-5. Reconnaissance d'un comportement anormal par raisonnement ontologique 169
Figure 4-6. Fréquentation d'une zone tampon autour du port de Toulon, pendant deux semaines . 170
Figure 4-7. Cas d'utilisation pour l'interface entre l'utilisateur et l'ontologie au sein d'un environnement d'aide à l'analyse géovisuelle 171
Figure 4-8. Architecture du module d'utilisation de la base de connaissances 172
Figure 4-9. Interface Java pour l'utilisation de la base de connaissances 174
Figure 4-10. Sélection du but d'utilisation *Analyse de comportement de pêche,* puis affichage des différentes tâches et données correspondantes 174
Figure 4-11. Fenêtre de choix des méthodes de visualisation après exécution du raisonneur dans l'ontologie 175
Figure 4-12. Accès à l'outil d'aide à l'analyse géovisuelle depuis FishEye : (a) par le menu général, (b) par un menu contextuel 176
Figure 4-13. Exécution de l'aide à l'analyse géovisuelle depuis FishEye, et utilisation du JNLP 177
Figure 4-14. Trajectoire d'un tanker (navire rouge) au Sud-Est de l'Espagne, et fiche descriptive de son identité 179
Figure 4-15. Classification Cynefin des évènements 180
Figure 4-16. Ecran de contrôle en centre de surveillance, les navires Paula-C et Darya Gayatri en navigation proche et parallèle, navire Raquel en manœuvre d'évitement 181
Figure 4-17. Trajectoires parallèles dans le pas de Calais, représentées après accident sur une carte électronique britannique 181
Figure 4-18. Ecran radar à bord du navire Paula C au moment de l'accident et affichage des données AIS 182
Figure 4-19. Représentation de trajectoires parallèles et graphe de vitesse correspondant 183
Figure 4-20. Représentation de trajectoires de navires parallèles par un cube spatio-temporel grâce (a) au logiciel R et (b) au logiciel ILWIS 184
Figure 4-21. Formes de trajectoires suspectes lors de l'approche d'un navire 185

Figure 4-22. Modélisation des processus d'analyse pour les cas (a) de dérive et (b) d'attaque pirate 189
Figure 4-23. Cas d'un contrôleur amené à tester une hypothèse de dérivé : prise en compte de son profil, du but, de la temporalité des données et des données disponibles 190
Figure 4-24. Résultats du raisonnement pour le cas d'un contrôleur et d'un scénario de dérive 191
Figure 4-25. Temps (en ms) utilisé par la raisonneur Pellet pour classifier les objets, selon le nombre de méthodes de visualisation modélisées 198

Liste des tables

Table 1-1. Bilan des étapes de la gestion des risques par la technologie 30
Table 1-2. Facteurs qualitatifs pour l'usabilité de l'information géographique (d'après Harding 2013) 50
Table 2-1. Variables visuelles et évaluation de la perception 68
Table 2-2. Variables de l'animation et dynamiques 77
Table 2-3. Typologie des données 87
Table 2-4. Typologie de l'information 88
Table 2-5. Modèle de Card et Mackinlay, agrémenté de l'extension dynamique 91
Table 2-6. Comparaison de moteurs d'inférence (d'après Dalwadi et al. 2012) 107
Table 3-1. Exemples de requêtes dans l'ontologie 119
Table 3-2. Exemple de questions d'exploration et d'analyse visuelle pour la prédiction de collision entre deux navires 130
Table 3-3. Description des caractéristiques retenues pour la description de l'utilisateur 134
Table 3-4. Exemples d'instances d'utilisateurs dans l'ontologie du domaine maritime 136
Table 4-1. Différenciation des scénarios identifiés, par la visualisation d'information a posteriori 187
Table 4-2. Ordre d'identification des évènements pour le scénario 1, Attaque de pirates 188
Table 4-3. Proposition de visualisations selon les utilisateurs et les scénarios étudiés 191
Table 4-4. Evènements reconnus par les systèmes actuels et par la GeoVA 193
Table 4-5. Scénario d'attaque, surveillé par un contrôleur maritime, et étapes de visualisation dans le temps 194
Table 4-6. Scénario d'attaque, étudié par un analyste 195
Table 4-7. Bilan du rôle de l'analyse géovisuelle dans un processus de gestion des risques 199

INTRODUCTION

Notre thèse s'intéresse à la visualisation du trafic maritime dans un contexte de sûreté et sécurité, relativement à la maritimisation de l'énergie. Dans cette introduction, nous présentons ce contexte international aux nombreux défis scientifiques, technologiques, sociologiques et politiques. Le mode de fonctionnement de la surveillance maritime en France est présenté, celui-ci étant à la base de notre problématique. Ainsi, nous pouvons mettre en évidence notre problématique de recherche, ainsi que les hypothèses et objectifs qui guident cette thèse.

Contexte : maritimisation de l'énergie

L'espace maritime, espace de rigueur, de liberté et de risques

Notre planète bleue tient son surnom de l'immensité de la surface terrestre couverte par des océans, mers, lacs et autres sources d'eau : plus de 71 % du globe. L'homme et la mer ont toujours eu un lien très étroit au cours de l'histoire de notre civilisation. Depuis l'Antiquité, elle représente l'origine de la vie sur Terre, mais aussi un milieu étranger à la nature humaine et terrestre. Les explorateurs, auteurs et poètes ont toujours vu en elle une source d'inspiration, mais aussi de crainte et d'émerveillement. En témoignent les nombreuses créatures représentées sur les anciennes cartes d'exploration maritimes, les portulans. Utilisées entre le XIII^{ème} et le XVIII^{ème} siècle, ces cartes de navigation repéraient les ports, ainsi que les dangers de navigation qui les entouraient : rochers affleurant, courants, etc. Afin de mettre au point de telles cartes, les navigateurs se basaient simplement sur leur propre perception de la côte maritime lors de leurs déplacements le long du littoral (cabotage), aidés de boussoles, de sextants et d'alidades. La carte et la mer étaient alors déjà étroitement liées pour l'exploration du monde, ainsi que pour la gestion des risques de navigation. Aujourd'hui encore, l'espace maritime reste un environnement de grand intérêt pour l'homme. Il joue en effet un rôle majeur dans les problématiques économiques, politiques et environnementales, notamment pour les problématiques énergétiques.

Chaque année, plus de 7,5 milliards de tonnes de marchandises sont transportées dans le monde par voie maritime. De 2010 à 2011, le trafic maritime mondial progresse ainsi de 4 %, pour atteindre un record historique de 8,7 milliards de tonnes, selon un rapport de l'ONU lors de la Conférence des Nations Unies sur le Commerce et le Développement (UNCTAD) (Division on Technology and Logistics 2013). Selon ce même rapport, la valeur de ces exports mondiaux transportés par voie maritime atteint 18,2 mille milliards de dollars en 2011. L'Europe est elle-même le premier pôle d'échange par voie maritime, représentant 23 % des échanges mondiaux en tonnage. Quant au nombre de passagers maritimes, recensé en 2010, il montre néanmoins une légère baisse depuis plusieurs années : lors de cette année, l'UE a par exemple compté 400 millions de voyageurs maritimes, soit une baisse de 2,2 % par rapport à 2009[1].

[1] Source Eurostat : appsso.eurostat.ec.europa.eu

Toutefois, l'espace maritime doit obéir à de nombreuses règles, qui contraignent son utilisation. Comme l'a si bien écrit Victor Hugo, *« la mer est un espace de rigueur et de liberté »*. Cette rigueur de l'espace maritime se traduit par une vaste réglementation nationale et internationale. Au niveau international, les sources de règlementation sont gérées par l'Organisation Maritime Internationale (OMI), créée en 1948. Un exemple notoire de collaboration internationale pour la gestion d'évènements majeur a été le naufrage de l'Erika (12 décembre 1999), un accident majeur de cette dernière décennie, qui a eu lieu dans les eaux internationales. Cette catastrophe a montré l'extrême nécessité d'une collaboration internationale au niveau de la règlementation maritime et des interventions des pouvoirs publics à mener lors d'un tel évènement (Muller et al. 2000). Les eaux internationales, ou haute mer, représentent les zones maritimes qui n'obéissent à l'autorité d'aucun Etat, soit 64 % des océans. Elles commencent à la limite des zones économiques exclusives (ZEE) des Etats côtiers, espace soumis au régime juridique établi par le pays (200 milles marins). Celui-ci concerne les droits souverains aux fins d'exploration, d'exploitation de gestion des ressources naturelles maritimes. De nombreuses autres zones règlementées pour le contrôle juridique de l'espace maritime ont été mises en place, au-delà des eaux territoriales et en complément des ZEE : les zones de pêche de la FAO, les zones de protection de pêche (ZPP), les zones de protection écologique (ZPE). Motivée par des intérêts écologiques et touristiques, la juridiction internationale reconnait aussi des aires marines protégées (AMP) pour la protection de la faune et la flore.

Mais la rigueur imposée par cette réglementation maritime n'empêche pas la persistance de nombreux risques, en haute mer tout comme sur les espaces côtiers. Malgré une baisse notable des attaques pirates le long des côtes somaliennes en 2013 (La Presse 2014), la piraterie reste une menace très importante sur les côtes de zones sensibles, notamment dans le golfe d'Aden, le golfe de Guinée (Slate 2014) et dans la mer de Chine méridionale (voir Figure 0-1). Des menaces telles que la piraterie portent atteinte à la **sûreté** de l'espace maritime, un aspect lié à la lutte contre le terrorisme contre les équipements et les équipages, ainsi que toute activité illicite. La **sécurité** maritime concerne plus généralement les navires, les personnes, les états et l'environnement. L'augmentation du trafic maritime mondial provoque en effet une densité importante de navires sur les routes maritimes ou dans les ports, augmentant ainsi les risques de collision, échouement, échouage (acte volontaire), etc. L'exemple notable du naufrage du Costa Concordia au large du littoral sud de la Toscane, le 13 janvier 2012, a profondément marqué les esprits. Bien que ce navire de croisière ait adopté un détour habituel, selon les dirigeants de la compagnie, aucune autorisation n'avait été donnée pour suivre cette route.

Au sein de ce contexte mondial, la mer joue un rôle majeur dans de nombreuses problématiques énergétiques : l'exploration, la production et le transport de l'énergie en mer sont autant d'aspects sensibles qui doivent être contrôlés. Pour cela, Napoli définit le nouveau concept de **maritimisation de l'énergie** (Napoli 2014), ainsi que les aspects de sécurité et sûreté relatifs à la maritimisation de l'énergie. L'étude de l'environnement maritime et du trafic des navires est un point fondamental pour anticiper et réduire les risques liés aux activités illicites, aux erreurs humaines et aux problèmes techniques qui menacent constamment cet espace (Davenport et Risley 2006 ; Adler et Fuller 2007 ; Glandrup 2013). Nous utilisons alors ce nouveau concept de **Sécurité et Sûreté de la Maritimisation de l'Energie (SSME)** que propose Napoli, afin de définir le contexte de notre recherche. Dans la sous-partie suivante, nous présentons les moyens mis en œuvre afin de contrôler les risques qui menacent la sécurité et la sûreté de l'espace maritime, soit la maritimisation de l'énergie de façon plus générale.

Figure 0-1. Infographie de la localisation des attaques pirates en 2011 (*source : Travelinsurance.org*)

Les acteurs de la « Sécurité et Sûreté de la Maritimisation de l'Energie »

Afin de faire face aux nombreux risques en mer qui ont été présentés précédemment, plusieurs moyens et acteurs entrent en jeu. Cela va de la mise en place de nouvelles règlementations, comme nous l'avons vu précédemment, gérées au niveau mondial par l'OMI. Dans la chaîne de la surveillance maritime, plusieurs organismes sont chargés de la gestion et de la surveillance du trafic maritime, de l'aide aux navires et personnels navigants à l'arrestation des pirates. Ces moyens mis en œuvre sont organisés de l'échelle du port (capitaineries) à l'échelle nationale. L'Action de l'État en mer (AEM) représente les moyens organisationnels mis en place sur le territoire maritime français, en matière de défense, protection et sauvegarde de l'environnement maritime. Dirigée par le Premier ministre français, ainsi que son représentant le préfet maritime, l'AEM comprend des administrations telles que la Marine nationale, la Gendarmerie nationale, les Affaires maritimes, les douanes, etc. La Figure 0-2 présente l'organisation de l'Action de l'État en mer et les interactions entre les différents acteurs.

Dans le cadre de l'AEM, des Centres régionaux opérationnels de surveillance et de sauvetage (CROSS) sont mis en place le long des côtes françaises pour assurer la surveillance d'espaces maritimes stratégiques. Les CROSS sont intégrés au réseau international des Centres de coordination de sauvetage maritime de la convention *Search and Rescue* (SAR) signée à Hambourg en 1979 (IMO 2014). En France métropolitaine, les CROSS sont au nombre de 5, se partageant ainsi les eaux françaises (voir carte en Figure 0-3) : Gris-Nez, Jobourg, Corsen, Étel et La Garde. Deux autres CROSS sont situés dans les DOM-TOM : Antilles-Guyane et La Réunion. Ces centres de surveillance du trafic maritime sont retrouvés dans tous les états côtiers, généralement sous le nom de MRCC, pour *Maritime Rescue Coordination Centre*. On retrouve ces utilisateurs dans de nombreuses publications

et rapports (Davenport et Risley 2006 ; van Laere et Nilsson 2009 ; Directorate-General for Maritime Affairs and Fisheries 2010 ; Lavigne et al. 2011 ; Glandrup 2013 ; Enguehard et al. 2013b)

En plus de ces centres de contrôle, des bureaux d'études spécialisés en risques maritimes ont pour rôle d'étudier des accidents passés, afin de comprendre le contexte de ces accidents et en tirer des leçons. De tels bureaux sont gérés par chaque pays côtier, tels que le BEAmer en France (Bureau enquêtes des accidents de mer) ou le MAIB au Royaume-Uni (*Marine Accident Investigation Branch*). Au sein de ces groupes, des analystes doivent étudier les données passées : ces informations gérées prennent en compte les manœuvres des navires accidentés, le contexte géographique, les échanges entre les navires et le centre de contrôle. Une description détaillée de l'historique des déplacements et de l'état des navires concernés est menés. En France, ce bureau d'étude est constitué au sein de l'Inspection Générale des Affaires Maritimes (IGAM) et placé directement auprès de l'inspecteur général des affaires maritimes. L'article 12 du décret du 26 janvier 2004 défini alors le rôle du BEAmer comme : réaliser des enquêtes techniques sur les événements de mer, en rechercher les causes et en tirer, sous forme de recommandations, les enseignements qu'ils comportent pour l'amélioration de la sécurité maritime. Le rôle de tels experts pour l'analyse et la modélisation des risques et comportements anormaux, pour l'amélioration de la gestion des risques maritimes est souligné dans plusieurs publications (Davenport et Risley 2006 ; Adler et Fuller 2007 ; Mansouri et al. 2009a ; Mansouri et al. 2009b ; Riveiro 2011 ; Riveiro et Falkman 2011).

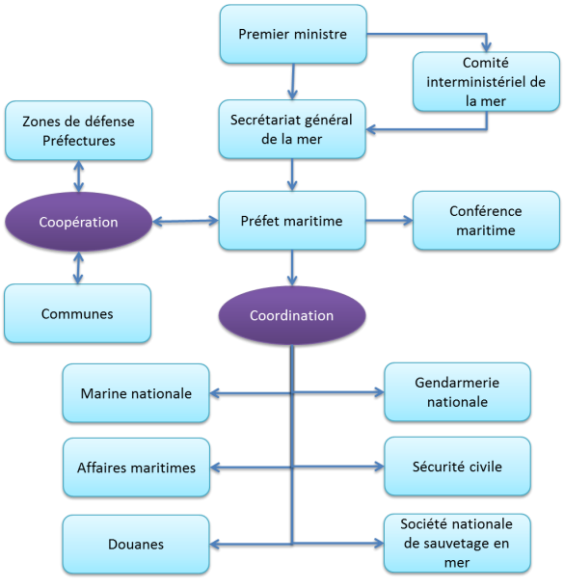


Figure 0-2. Organisation de l'*Action de l'Etat en mer*, en France

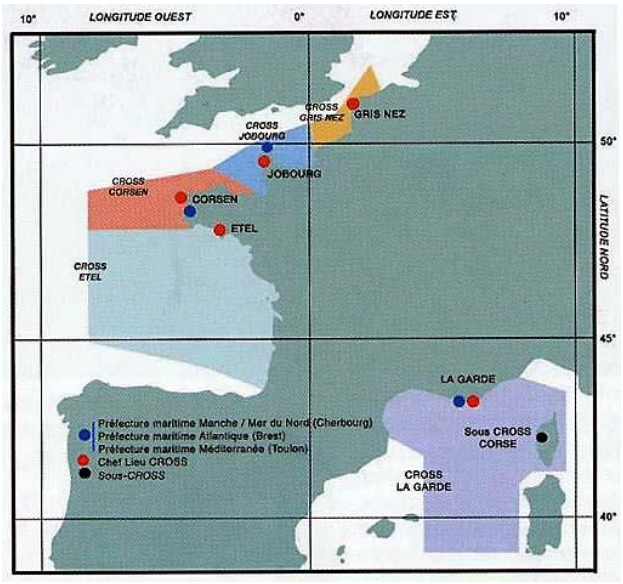


Figure 0-3. Répartition des CROSS en France métropolitaine (*source : marine-marchande.net*)

Des organismes spécialisés selon le type de menace ont aussi récemment été mis en place par les gouvernements, avec leur propre équipe de reconnaissance et d'action. Par exemple, le Centre de Coordination de la Lutte Anti-drogue en Méditerranée (CeCLAD-M), basé à Toulon, permet de lutter contre le trafic illicite de drogue, grâce à une étroite collaboration avec des services européens et mondiaux de renseignement. Un autre exemple est le système EUROSUR, spécialisé dans la surveillance des frontières extérieures de l'espace Schengen. Cet outil permet des échanges d'informations et la collaboration des forces européennes pour la lutte contre la criminalité transfrontalière, l'immigration clandestine, ainsi que la protection et le sauvetage des migrants en mer. Cet outil très récent a été mis en place en décembre 2013.

Ces nombreux acteurs de l'AEM et de la surveillance maritime européenne collaborent pour la défense et la protection de l'espace maritime international, et doivent utiliser de nombreux outils et bases de données pour mener à bien leurs missions. Dans la prochaine partie, nous présentons les systèmes utilisés pour la surveillance du trafic maritime, ainsi que leurs perspectives d'évolution.

Les outils pour la navigation et la surveillance maritime

Au sein de cette chaîne de surveillance complexe, les informations de mobilité transmises de manière active ou passive par les navires servent de données de base pour la surveillance des déplacements des navires. En effet, l'OMI a mis en place l'obligation d'utiliser des transpondeurs AIS pour une certaine catégorie de navires : les navires de tonnage brut supérieur à 300 en déplacement international, les cargos de tonnage brut de plus 500 en déplacement non international, ainsi que tous les navires à passagers[2]. Les informations AIS sont basées sur des échanges de messages par radio VHF, comportant des informations sur la nature du navire, son chargement, sa position et sa dynamique. Ces informations proviennent d'outils de mesure tels que les GPS / DGPS, les compas gyroscopiques et les lochs. Ces messages sont envoyés de manière régulière par les navires en déplacement, dont la fréquence dépend de leur vitesse. Néanmoins, il est important de noter les limites de la qualité de ces données, une partie étant renseignée manuellement par le chef de quart. Ces AIS sont souvent couplées à des informations Radar pour compléter et améliorer la qualité des informations reçues. Nous pouvons envisager que dans le futur, l'obligation d'utiliser les transpondeurs AIS pourra être élargie à un nombre beaucoup plus important de types de navires. Par exemple, les navires de plaisance ou les navires à grande vitesse, qui ne sont pas encore soumis à ce type de règlementation, mais sont pourtant les victimes ou les causes de nombreux accidents.

Figure 0-4. Ecran radar pour l'aide à la navigation maritime (*source : Wikimedia*)

Historiquement, les radars ont été la première source d'information pour la surveillance du trafic maritime ou aérien. En effet, l'application première du radar maritime a été de détecter les obstacles en mer, notamment avec les radars de navigation ou les radars anti-collision (Figure 0-4). Dans la zone de portée des radars, cette source d'information est complémentaire aux données AIS et permet alors de détecter les navires qui ne sont pas soumis à l'obligation d'être équipé de

[2] www.imo.org/OurWork/Safety/Navigation/Pages/AIS.aspx

transpondeur AIS. Inversement, les AIS permettent de compléter les informations Radar, lorsque les ondes radio ne peuvent pas circuler dans un certain contexte géographique (présence de masque par le terrain).

Ces nombreuses sources de données permettent alors de suivre en temps réel le trafic maritime, que ce soit depuis un centre de surveillance (systèmes de surveillance maritime, VMS, VTS) ou bien depuis les navires (systèmes d'aide à la navigation, systèmes anti-collision). En plus des données de mouvement des navires, des données cartographiques sont nécessaires afin de connaitre le contexte géographique, topographique, règlementaire, etc. Pour cela, les **cartes électroniques de navigation** vectorielles (ENC) constituent une base de données cartographique pour l'affichage des informations maritimes (bathymétrie, zones règlementées, etc.), gérées par l'OHI. Ces cartes sont basées sur les **normes S-57** pour l'enregistrement et l'échange des données (IHO 2000), ainsi que les **normes S-52** pour leur affichage (IHO 2010). Les ENC peuvent être visualisées grâce aux systèmes de visualisation des cartes électroniques et d'information, appelés ECDIS (*Electronic Chart Display and Information System*). Ceux-ci doivent suivre les normes mises en place par l'OMI sur les bases de données maritimes et leur représentation (ISO 2003, 19379) : *Navires et technologie maritime - Bases de données ECS - Contenu, qualité, mise à jour et essais*. Les ECDIS fournissent des informations de positions des navires proches, les dangers, les éphémérides et la signalisation maritime. Malheureusement, ces systèmes sont régulièrement soumis à de trop nombreuses données à afficher, ce qui rend leur lecture et leur interprétation compliquée, voire impossible (Figure 0-5). De plus, ces outils d'affichage ne disposent pas de fonctionnalités d'analyse du trafic (Glandrup 2013).

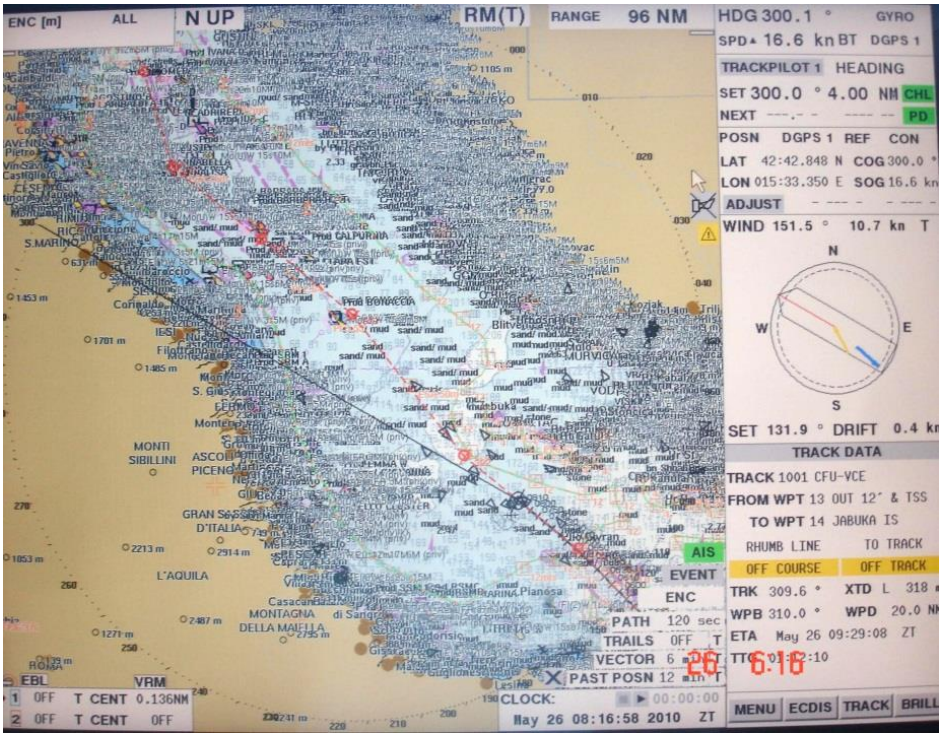


Figure 0-5. Surcharge d'information sur un ECDIS (*source : nautinst.org*)

Des outils de suivi du trafic maritime ont pu être développés grâce aux technologies web et ouverts au public, en se basant l'affichage des données AIS. Nous pouvons citer, entre autres, les sites Marine

Traffic[3], ShipFinder[4] ou GeoGarage[5] qui recensent toutes les informations des bases de données AIS Hub[6]. Il est important de noter que ces plateformes cartographiques ne permettent pas d'effectuer une **analyse des mouvements** des navires, mais seulement de suivre leurs déplacements : il n'y a pas d'interprétation des comportements, en vue de détection d'anomalies, de dangers, etc. Afin de mener ces opérations de surveillance et d'analyse des menaces et dangers en mer, rôle premier de l'Action de l'État en mer, il est nécessaire de mettre en place des technologies et outils pour analyser et interpréter les mouvements des navires.

Analyse et détection des comportements à risques : les systèmes de surveillance de nouvelle génération

La surveillance maritime concerne le contrôle du trafic maritime en pleine mer ou à l'échelle des ports, dans le but de détecter des comportements anormaux ou dangereux. Ces comportements peuvent causer de graves accidents, tels que des collisions entre navires, la pollution de l'environnement maritime, ou même menacer les passagers ou l'équipage des navires et autres installations en mer. Dans le travail de surveillance, les systèmes de surveillance maritime (SSM) permettent de suivre en quasi temps réel le trafic dans une zone donnée. Un SSM nécessite une utilisation de nombreux écrans de contrôle, de manière à donner simultanément un aperçu cartographique des positions des navires, la description des navires affichés, une liste des actions en cours, etc.

Les SSM utilisés actuellement par les opérationnels permettent le suivi dynamique des navires. Par exemple, Trafic 2000 et sa version allégée SafeSeaNet permettent des échanges d'informations sur le trafic et la conformité des navires quant aux normes de sécurité. Le système SPATIONAV utilise les données de radars et sémaphores pour échanger des informations entre les administrations de l'AEM, et propose des outils d'aide à la décision (extrapolation, rejeu de situation, évaluation des conditions d'intervention, alarmes opérationnelles) ainsi que des outils de communication entre les acteurs. En 2013, la version SPATIONAV V2 a commencé son déploiement, afin de centraliser ces informations avec l'Union Européenne. D'autres outils sont des bases de données dressant un historique des situations, tels que la base SIRENAC qui enregistre les contrôles des navires.

Néanmoins, les systèmes de surveillance utilisés aujourd'hui de façon opérationnelle posent de nombreuses limites, notamment sur le traitement et l'analyse des informations (Roy et Davenport 2009 ; Glandrup 2013). Le nombre très important de navires à surveiller, la multiplicité des risques et des règles, ainsi qu'un possible laxisme dans l'application des mesures de sécurité (accident du Costa Concordia), sont toutes autant de raisons qui demandent de nouvelles infrastructures en matière de sécurité et sûreté maritime (European Commission 2008 ; Royal Institution of Naval Architects 2014).

Depuis une dizaine d'années, de nombreux projets R&D nationaux et internationaux contribuent au développement de nouvelles solutions d'**intégration**, de **suivi** et d'**analyse automatisée** des données de trafic maritime, ainsi que l'**amélioration des capteurs** de données (radars FM-CW, caméras, etc.). La Figure 0-6 présente un schéma des différentes étapes et modules pour l'intégration de données et

[3] http://www.marinetraffic.com
[4] http://shipfinder.co
[5] http://carte-france.geogarage.com
[6] http://www.aishub.net

connaissances pour améliorer les systèmes de surveillance maritimes. Les travaux de recherche menés sur les nouveaux capteurs maritimes, aériens ou terrestres se situent dans la première partie de ce schéma (*Surveillance*). Les nombreux travaux de recherches actuels se situent majoritairement dans la partie centrale, la *Détection* des anomalies. Nous pouvons citer par exemple les nombreux travaux français et européens coordonnés par les groupes DCNS ou Thales :

- ***Pirana*** (2006), pour une antenne radar à émission-réception distribuée en bande X à bas coût pour la surveillance maritime aéroportée
- ***ScanMaris*** (2007) et Tamaris (2008), des projets ANR pour le traitement, l'authentification des menaces en mer (Morel et al. 2010 ; Morel et al. 2011)
- ***Sismaris*** (2008), un projet ANR pour l'intégration de ces résultats dans un système d'informations (Morel 2009)
- ***Sargos*** (2010), un projet européen pour une réponse graduée aux attaques de pirates sur les plateformes pétrolières offshore (Chaze et al. 2012)
- ***I2C*** (2010), un projet européen pour l'identification et le suivi de comportements à risques à l'aide de capteurs innovants (Morel et Claisse 2010)
- ***Reconsurve*** (2010), un projet européen d'analyse du trafic maritime basé sur des règles reconfigurables et de nouveaux capteurs (Ray et al. 2013)

Le contrôle de la sécurité (des personnes, prévention de la pollution) et de la sûreté (prévention et lutte contre les actes illicites) demande une exploration approfondie et une compréhension précise des profils des objets mobiles et de leur contexte. La **définition** (modélisation) des comportements anormaux ou à risques est fondamentale, afin de les détecter par des méthodes automatiques ou bien visuelles. Cette tâche représente notamment le **rôle des chercheurs** dans le cadre de la sécurité et sûreté maritime. Dans cette optique, les SSM prennent part à la gestion des risques appliquée au trafic maritime, notamment dans la reconnaissance des anomalies et des menaces au sein de données en temps réel ou de données passées.

Comme nous allons le voir par la suite, la gestion de risques s'appuie à la fois sur une surveillance en **temps réel** du système d'intérêt, mais aussi sur l'**analyse de données passées** (accidents, routes de navires, etc.). Dans le cas de données passées, l'analyse a posteriori a pour but de rechercher des comportements qui sortent de la normale parmi de grandes quantités de données de mouvement, afin de caractériser ces menaces. Cette caractérisation peut être utilisée ensuite pour l'analyse des données en temps réel, afin de les comparer aux prémices connus de comportements à risques. Les travaux menés au sein de notre laboratoire se sont fortement appuyés sur cet axe de recherche depuis cinq ans, et prennent part au module final présenté sur la Figure 0-6 : l'**analyse des informations maritimes**. En effet, les travaux de recherche de notre centre ont mené à la proposition de nouvelles méthodes qui permettent de :

- **Découvrir** des connaissances dans de grandes bases de données accidents et déplacement, grâce à des techniques de fouille de données (Idiri 2013 ; Idiri et Napoli 2012a)
- **Modéliser** les évènements et des risques maritimes, à l'aide d'ontologies intégrant les notions topologiques (Vandecasteele 2012 ; Vandecasteele et al. 2014)
- **Guider** dans l'analyse de l'information géographique pour mieux découvrir et comprendre les risques maritimes (Vatin et Napoli 2013c ; Vatin et al. 2014)

- **Anticiper** la réaction face à des attaques pirates, avec une approche probabiliste, par l'utilisation de réseaux bayésiens (Bouejla et al. 2012 ; Chaze et al. 2012 ; Bouejla et al. 2014)

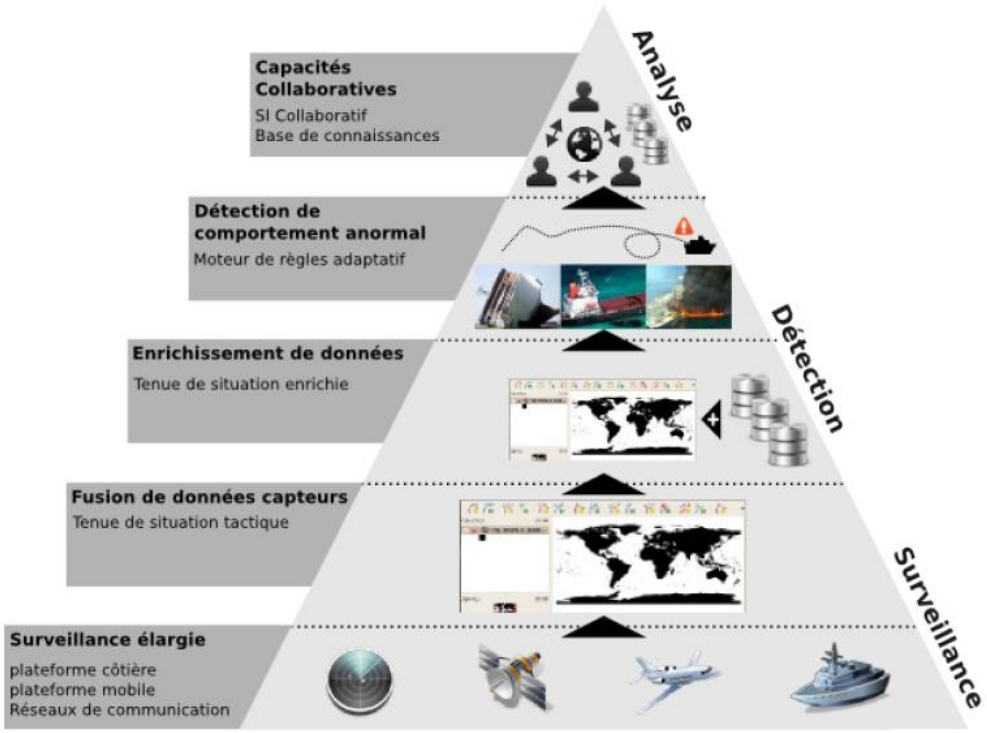


Figure 0-6. Amélioration des composantes des systèmes de surveillance (*source :* Vandecasteele 2012)

Dans ces publications récentes, nous pouvons voir que les recherches menées en amont explorent de nouveaux axes afin d'améliorer la surveillance maritime. Ces recherches ont exclusivement porté sur des processus automatisés, ou semi-automatisés, pour l'analyse des données d'objets mobiles. Néanmoins, comme nous allons le voir par la suite, le rôle de d'utilisateur humain de ces systèmes n'est pas à négliger. Autrement, de nombreuses limites liées à l'automatisation peuvent provoquer un effet de « boite noire » qui masque le processus de construction de connaissances (comme dans la fouille de données), un nombre trop important d'alertes, ou encore la perte de confiance et d'identité professionnelle de la part de l'utilisateur.

En s'inscrivant dans la continuité de ces travaux pour proposer des améliorations au contrôle et à l'analyse du trafic maritime, nous estimons qu'il est essentiel de revenir à la question principale dans le domaine sécurité et sûreté de la maritimisation de l'énergie : la « surveillance et la vigilance contextuelle » (*Situation Awareness*) grâce à des outils visuels. Nous estimons que, parallèlement aux recherches sur une automatisation du contrôle des mouvements, la manière de visualiser l'information par les outils de surveillance doit être améliorée, afin que l'homme, au cœur de la décision, puisse lui-même comprendre les données de trafic et prendre des décisions adaptées. Pour cela, notre recherche s'intéresse particulièrement aux apports de l'analyse géovisuelle pour la compréhension des données de mouvement. L'analyse géovisuelle, qui se base sur les principes de l'analyse visuelle et de la géovisualisation, permet en effet d'améliorer la compréhension de données à composantes spatiales, en améliorant la *capacité à extraire et organiser des informations émanant de l'environnement visuel* (Schneck 2005). Dans la partie suivante, nous présentons notre axe d'étude dans ce domaine de recherche.

Problématique de recherche : aide à l'analyse géovisuelle des données de mouvement

La partie précédente a introduit les nombreuses problématiques de l'espace maritime, ainsi que les travaux effectués pour l'amélioration de la sécurité et sûreté maritime. Nous avons notamment identifié l'apport de nouveaux systèmes pour l'analyse automatisée des comportements des navires, c'est-à-dire leur dynamique, mise en relation avec le contexte géographique où ils évoluent (topographie, dangerosité, etc.). Toutefois, la visualisation de ces informations ne doit pas être négligée, car il s'agit de l'interaction principale que l'humain (contrôleur, analyste, etc.) peut avoir avoir avec le système contrôlé (l'ensemble des navires en mouvement).

La visualisation, une approche anthropocentrique pour la prise de décision

Bien que les outils cartographiques utilisés actuellement dans les centres de surveillance maritime permettent d'assurer un contrôle suffisant pour la sécurité des navires, il est nécessaire de prendre du recul par rapport à leur utilisation. Les opérateurs et les analystes se basent fortement sur leur propre expérience et leur connaissance du « terrain », et obéissent à certaines habitudes afin de pouvoir détecter visuellement des anomalies au sein de données en temps réel. En effet, les opérateurs de surveillance fondent principalement leurs décisions sur leur propre interprétation des mouvements des navires, par la représentation sur des affichages ECDIS.

Néanmoins, les systèmes de surveillance présentés précédemment (Spationav, SafeSeaNet, etc.) ne proposent qu'un affichage des informations de trafic, et non pas des outils pour l'analyse du trafic (Glandrup 2013). Cette situation se retrouve aussi bien dans les centres de contrôle pour le trafic aérien, routier, etc. Comme nous pouvons l'observer dans les nombreux projets menés en R&D, une assistance automatisée ou semi-automatisée dans cette exploration des données de mouvement est alors proposée. Celle-ci permet la découverte de comportements jusqu'alors non connus, reconnus, et d'avoir ainsi une connaissance plus approfondie des comportements à risques.

De plus, le développement de nouveaux capteurs (radars FM-CW, drones, AIS) pour l'acquisition des données maritimes menacent de compliquer la surveillance du trafic maritime, en augmentant de manière exponentielle le nombre de données d'objets mobiles à surveiller (petites embarcations, jet skis, etc.). La charge cognitive des tâches visuelles d'analyse des données de trafic va donc être considérablement alourdie dans les années à venir.

L'optimisation des systèmes pour la surveillance et l'analyse des données de trafic passe alors par deux grands types de propositions, comme nous pouvons l'observer actuellement dans l'état de la recherche :

- L'automatisation de la reconnaissance des risques et menaces dans les données temps réel (Mansouri et al. 2009a ; Adler et Fuller 2007 ; Malik et al. 2012 ; Enguehard et al. 2013b). Celle-ci a notamment été identifiée au sein de projets R&D proposés depuis une dizaine d'année, et parmi lesquels le centre de recherches sur les Risques et les Crises a pu être amené à participer (Morel 2009 ; Morel et al. 2011 ; Etienne 2011 ; Ray et al. 2013 ; Idiri 2013 ; Bouejla et al. 2014).

- L'optimisation de la représentation cartographique des informations et l'aide à l'analyse par la visualisation (Riveiro et Falkman 2011 ; Willems et al. 2011 ; Mallé-Noyon 2008 ; Devogele 2009).

L'utilisation d'algorithmes automatisés pour la détection de comportements d'intérêt, qui correspond au premier type de proposition, nécessite une longue phase de modélisation des risques au préalable (Riveiro et Falkman 2011 ; Idiri 2013). De cette modélisation dépendent fortement la qualité des alertes émises et les décisions qui sont prises. De plus, il est très important de savoir qu'un environnement d'aide à la décision trop automatisé peut poser des limites quant à la justification de ces décisions et des actions menées sur le terrain (Enguehard et al. 2013a). Dans le cas d'analyse des mouvements d'objets mobiles tels que les navires et de la prise de décisions dans un cadre règlementaire (acte illicite), la compréhension des anomalies et la justification sont fondamentales. Le rôle de l'être humain dans l'analyse et la prise de décisions relatives à la sécurité et la sûreté ne peut donc pas être entièrement remplacé par un processus automatisé.

Dans cette optique, il est donc essentiel de mettre en place des environnements d'exploration basés sur la visualisation d'information qui soient adaptés aux tâches de surveillance et d'analyse, afin d'améliorer les processus cognitifs d'exploration menés par l'humain. De tels environnements de visualisation et d'analyse des données permettent en effet d'anticiper l'analyse de trop grandes quantités d'information, notamment pour des données d'objets mobiles (Andrienko et Andrienko 2013 ; Demšar 2009). De cette manière, la détection des comportements n'est pas automatisée, mais les environnements d'analyse géovisuelle peuvent aider l'utilisateur à **construire un raisonnement d'exploration et d'analyse** des données de mouvement, tels que les déplacements des navires.

Pour cela, nous pouvons qualifier cette approche d'**anthropocentrique**, comme Aristote a pu définir ce concept philosophique : il place l'homme comme *entité centrale la plus significative de l'Univers, et appréhende la réalité à travers la seule perspective humaine*. Une approche anthropocentrique de l'analyse d'information et de la prise de décision est donc une approche disjointe de l'analyse automatisée de l'information et des prises de décisions, classification, etc., qui sont basées uniquement sur un raisonnement machine automatisé.

L'espace géographique, support de l'analyse des mouvements

Les objets mobiles sont, avant tout, définis par leurs propriétés spatiales et temporelles telles que leur vitesse, leur direction, leur trace passée, etc. Le système composé d'un ensemble d'objets mobiles est donc défini dans un certain espace géographique et sur une période de temps donnée. Le principal but de la surveillance et l'analyse des comportements de ces objets mobiles est de **comprendre leur organisation spatiale** et leur fonctionnement, afin d'identifier des zones d'intérêt (dangers, haute densité, etc.), de **comprendre le contexte d'un évènement** et d'**anticiper les comportements futurs** et potentiellement dangereux (Glandrup 2013 ; Hollnagel et al. 2010 ; Davenport et Risley 2006). Pour cela, l'espace géographique joue un rôle majeur dans l'interprétation des données de mouvement, par la mise en contexte des données.

Voiron et Chéry (Voiron et Chéry 2005) soulignent l'importance de distinguer les termes d'espace et d'étendue : selon les auteurs, *l'espace géographique est un tout complexe formé d'éléments visibles, qui interagissent entre eux et constituent sa structure*. L'espace géographique est donc en constante évolution, de par les interactions entre les agents qui le constituent, et la façon dont ils s'adaptent

continuellement à leur contexte géographique : c'est l'aspect **organisé et organisant** de l'espace géographique mis en avant par les auteurs (Voiron et Chéry 2005, 10). Belhedi (Belhedi 1993) va jusqu'à mettre en avant une triple complémentarité des dimensions qui définissent cet espace : une dimension mathématique / géométrique, une dimension historique et une dimension socio-cognitive.

L'évolution des cartes dont l'histoire a pu être le témoin montre aujourd'hui encore l'évolution du rapport de l'homme à l'espace géographique : espace de l'imaginaire, du mythique et du religieux (cartes en T), espace de possibilités pour le commerce (portulans pour la navigation), espace social (cartographie d'indicateurs sociaux), espace de vulnérabilité (cartographie des risques), espace de possibilité (design urbain, SIG), et aujourd'hui espace de partage (OpenStreetMap) voir de médiation, selon Belhedi (Belhedi 1993). La Figure 0-7 donne un frise chronologique des évènements marquants du Moyen-Âge et de l'évolution de la cartographie publiée à cette période.

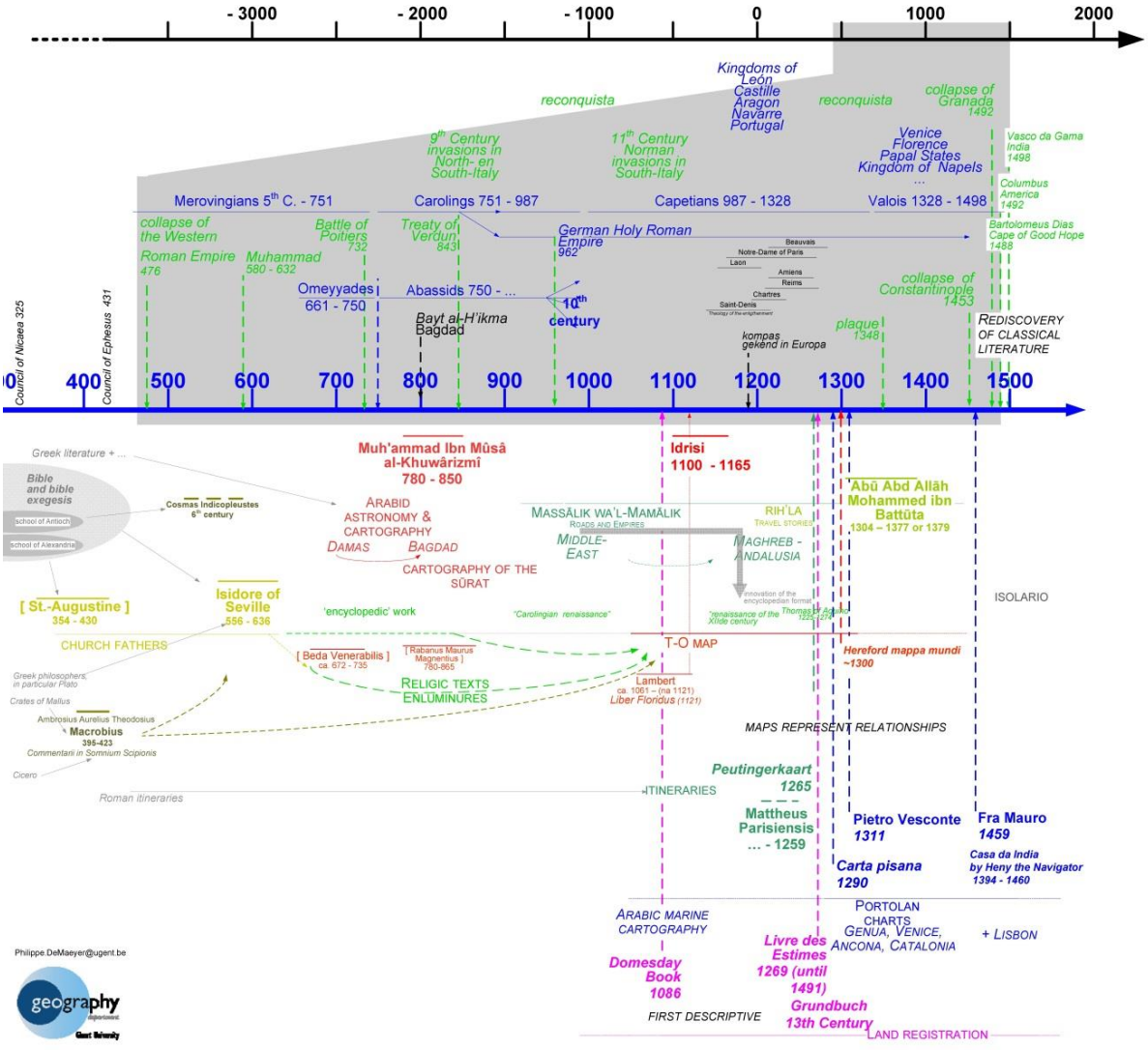


Figure 0-7. Evolution de la cartographie pendant le Moyen-Âge (*www.bimcc.org*)

Lorsque nous parlons de gestion de crises ou de risques dans un certain domaine spatial (la mer, le ciel, la route…), contrairement aux risques en entreprises, l'espace géographique est au cœur de toute l'analyse et de la prise de décision. Tout comme la physique quantique peut expliquer de très

nombreuses interactions entre des champs de forces, l'espace géographique peut expliquer des liens nombreux entre tous les agents qui le constituent.

Aujourd'hui, de très nombreuses technologies nous permettent d'acquérir des informations de mouvement en quasi-temps réel et d'une précision de plus en plus fine. L'espace géographique est alors la base de l'étude de ces données et de leur interprétation, notamment par la visualisation. Ce moyen est en effet le premier moyen de communication et d'analyse des évènements liés à l'espace géographique, et le plus simple à comprendre pour un utilisateur humain. L'espace géographique utilisé dans un but d'analyse d'information a donc demandé d'introduire, il y a une vingtaine d'année, un nouveau concept pour l'utilisation de l'information : la géovisualisation.

Une approche formelle pour la géovisualisation

La géovisualisation, qui peut être définie comme **l'utilisation d'interfaces visuelles interactives pour l'exploration de données à composantes spatiales** (MacEachren 1995), reprend de nombreux principes qui ont évolué avec les domaines de la visualisation d'information (InfoVis) et de la visualisation scientifique (SciVis). Pour cela, de très nombreuses méthodes permettant de visualiser des informations abstraites ou concrètes (i.e. information scientifique) sont reprises en plus des méthodes de visualisation d'informations purement géographiques. L'interaction entre ces nombreuses dimensions de la visualisation permet de prendre en compte de nombreux types de données et les relier sur différents plans (temporel, attributaire, géographique, etc.) (Peuquet 1994 ; Kraak 2006 ; Robinson 2011).

Ces très nombreuses possibilités de visualiser l'information peuvent poser problème, lorsque l'on doit comparer et étudier les apports de chacune de ces méthodes pour un problème donné. Pour cela, il est donc nécessaire de décrire ces nombreuses visualisations par des moyens formels, tels que des modélisations de ces méthodes. De plus, utiliser une approche formelle pour les caractériser permet de partager et réutiliser les connaissances issues de ces modèles (Fabrikant 2001) : à quoi servent ces méthodes, que peut-on faire ou ne pas faire ? Comme nous le verrons dans le chapitre 2 de ce manuscrit, de nombreux modèles ont pu être développé pour caractériser les méthodes de visualisation, qui sont alors caractérisées comme des *projections de l'information de base vers un espace visuel* (Fabrikant et Skupin 2005).

Vers un environnement d'aide à l'analyse géovisuelle

Cette dernière décennie, les laboratoires de R&D ainsi que les grands groupes liés aux sciences de l'information géographique (GIScience) proposent et développent de nouvelles manières de cartographier l'espace et visualiser l'information spatio-temporelle. L'utilisation d'environnements interactifs est la base de ces développements, permettant ainsi à l'utilisateur de contrôler son exploration des données, la manière de les afficher (légende, animation, etc.). Un tel engouement de la communauté GIScience pour ces nouvelles méthodes d'exploration, d'analyse et de visualisation de l'information à caractère géographique pose néanmoins un grand inconvénient. La multitude des outils et environnements pour visualiser l'information et la diversité des complexités proposées (nombreux paramètres, 3 dimensions, réalité augmentée ou virtuelle, etc.) rend leur utilisation beaucoup plus complexe qu'elle ne devrait l'être. Devant une trop grande panoplie d'outils disponibles, il est aujourd'hui compliqué de savoir quelle méthode, quel algorithme, quel logiciel utiliser pour mener un processus d'analyse d'information à composantes spatiales.

L'un des grands défis posés par la communauté GIScience ces dernières années n'est plus de proposer de nouvelles méthodes de visualisation de l'information, mais consolider l'utilisation de la visualisation, en étudiant le réel apport de ces nombreuses méthodes **par rapport aux questions de l'utilisateur**. Pour notre recherche, nous introduisons donc le concept de **contexte d'utilisation** de l'information géographique, qui représente la situation dans laquelle se trouve l'utilisateur :

- Les caractéristiques propres au profil de l'utilisateur
- Le contexte géographique d'utilisation : lieu, mobilité, etc.
- Le contexte technologique : ressources matérielles et logicielles disponibles
- Les données disponibles
- Les questions posées par l'utilisateur (tâches d'exploration)

Le choix d'une méthode de visualisation d'information dépend de ce contexte d'utilisation (Endsley et al. 2003 ; Lloyd et al. 2007). Les recherches passées dans ce domaine se sont surtout intéressées à la caractérisation des données, du profil de l'utilisateur et de son environnement géographique et technologique. Mais peu de travaux ont été faits dans la concordance entre les méthodes de visualisation et les tâches de l'utilisateur (Fabrikant 2001).

Hypothèses et objectifs

Afin de guider cette recherche, nous posons plusieurs postulats sur lesquels nous reviendrons avec plus de précisions dans les deux premiers chapitres de ce manuscrit. Ceux-ci posent une base reconnue par la communauté, qui justifie alors notre méthode de recherche.

Postulat 1 : Garder le contrôle par l'humain. Par habitude, les acteurs de la sécurité maritime amenés à étudier les données de trafic utilisent des outils visuels (les systèmes de surveillance maritime, les systèmes de suivi de navires, etc.). Bien que des algorithmes et des outils puissants permettent d'automatiser ces tâches d'exploration et d'analyse de données pour améliorer le contrôle de la situation, l'utilisation d'outils visuels est toujours favorisée par les opérationnels et les décideurs (Noyes et Bransby 2001 ; Davenport et Risley 2006 ; Riveiro et Falkman 2011). De cette manière, les évènements observés sont mieux compris et mieux retenus que si ces résultats résultaient d'un processus entièrement et seulement automatisé. De plus, le rôle de l'expertise humaine est fondamental dans un contexte de prise de décision, où des vies humaines peuvent être en jeu.

Postulat 2 : La géovisualisation comme technique d'analyse des comportements d'objets mobiles. De très nombreuses solutions d'analyse géovisuelle ont pu être mises au point par les communautés scientifiques de la visualisation d'information, et des sciences de l'information géographique (Kraak 2006 ; Andrienko et Andrienko 2013). La géovisualisation et l'analyse géovisuelle sont des domaines de recherche présentant de grandes avancées dans le développement d'interfaces homme-machine (IHM) pour l'exploration et l'analyse de données spatio-temporelles ; notamment les données de trafic de tous types d'espaces (route, ciel, mer, etc.) (Hurter 2010 ; Guo et al. 2011 ; Willems 2011 ; Riveiro 2011). L'analyse géovisuelle permet de découvrir de nouvelles connaissances grâce à une exploration visuelle des données, tout en évitant une surcharge cognitive.

En se basant sur ces deux constats, l'utilisation de l'analyse géovisuelle pour l'étude des données de trafic montre tout son intérêt : de nombreuses méthodes d'analyse géovisuelle ont été développées

par la communauté scientifique en GeoVis afin d'étudier la dynamique et le contexte de données d'objets mobiles et afin de leur donner du sens. Cette exploration étant menée par la visualisation, l'humain est au cœur de cette démarche d'extraction de connaissances. Il gère ainsi de manière plus compréhensible les connaissances issues de ce processus, contrairement à des processus automatisés qui demandent de nombreuses vérifications et interprétations des résultats.

Néanmoins, la multitude des solutions d'analyse géovisuelle pose un grand problème quant à leur utilisation : comment connaitre la solution la plus adaptée au contexte d'utilisation. Certaines grandes questions peuvent donc se poser, lorsque l'on constate la diversité des solutions proposées par la communauté GeoVis : quand utiliser une méthode plutôt qu'une autre ? Qui peut l'utiliser ? A quelle situation appliquer un certain type de visualisation ?

Le point fondamental n'est donc pas de développer une nouvelle méthode pour l'analyse géovisuelle des trajectoires, mais de rechercher comment guider l'utilisateur dans le processus d'analyse géovisuelle. Il est en effet nécessaire de faire correspondre des solutions existantes au contexte d'utilisation (tâches, utilisateur, données, environnement de travail, etc.) dans le processus d'analyse des comportements à risques. Afin de guider notre travail de recherche, nous posons plusieurs hypothèses par rapport à ce questionnement. Celles-ci seront démontrées durant notre recherche, notamment par le développement d'un outil d'aide à l'analyse et l'évaluation de visualisations.

Hypothèse 1 : Assister l'opérateur dans l'utilisation de solutions géovisuelle permet d'améliorer l'analyse des situations à risques. L'analyse géovisuelle permet d'étudier des évènements spatio-temporels. Les informations sur la géométrie, la dynamique et le contexte des objets sont comparés, permettant ainsi une meilleure analyse et compréhension des risques et accidents. Si l'utilisateur dispose d'un guide dans l'utilisation de méthodes adaptées à son environnement et ses besoins, nous supposons que son travail d'analyse des risques doit en être amélioré.

Hypothèse 2 : La formalisation du contexte d'utilisation et des solutions de visualisation permet d'aider l'utilisateur dans son processus d'analyse géovisuelle. La modélisation des utilisateurs, des utilisations et de solutions d'analyse géovisuelle peut être représentée par un langage formel et partageable. Ce modèle doit avoir pour but de formaliser les tâches à accomplir dans le cas de l'analyse des trajectoires d'objets mobiles dans un contexte de surveillance, ainsi que de formaliser les possibilités offerte par les méthodes d'analyse géovisuelle. Ce modèle doit mettre en évidence les interactions et les visualisations les mieux adaptées à ces tâches identifiées. Un outil basé sur cette base de connaissance doit permettre d'assister dans l'utilisation et/ou la conception d'environnements d'analyse géovisuelle adaptés au contexte.

Objectif de la thèse : Formalisation d'un « environnement d'aide à l'analyse géovisuelle », et application à la sécurité et sûreté de la maritimisation de l'énergie.

Au cours de notre recherche, l'étude du contexte de la surveillance maritime, puis la généralisation à l'analyse d'objets mobiles, un élément clé a pu fixer notre objectif principal de recherche. Cette thèse a pour but de définir ce que nous avons appelé un *environnement d'aide à l'analyse géovisuelle* : nous définirons un tel environnement dans le chapitre 3 de cette thèse. L'objectif de cette thèse est

de montrer en quoi ce concept nous a semblé être un point de recherche fondamental, et comment il pourrait être développé. Son application au domaine de l'analyse des comportements des navires en mer vient donc appuyer cette démarche. Au sein de cet objectif principal, nous pouvons donc identifier les étapes suivantes dans notre travail de recherche :

- Etude de l'utilisation actuelle des systèmes de surveillance maritime
- Etude de l'usage de l'analyse géovisuelle pour l'analyse de comportements d'objets mobiles
- Définition d'un environnement d'aide à l'analyse géovisuelle
- Modélisation d'un tel environnement et du processus d'analyse de comportements
- Développement d'un prototype d'aide à l'analyse géovisuelle

Démarche de recherche

Suite à de précédentes thèses, une recherche bibliographique a été menée en parallèle dans les domaines de la sécurité maritime et de géovisualisation. Cette étape nous a permis d'identifier des limites d'utilisation des outils de surveillance actuels, ainsi que d'identifier le domaine de l'analyse géovisuelle comme « solution » à de nombreux problèmes d'analyse de l'espace maritime. La recherche bibliographique dans ce domaine récent nous a aussi permis d'identifier certaines limites quant à l'utilisation des nombreuses méthodes d'analyse géovisuelle : de cette manière, une problématique commune à ces deux communautés nous ont permis de définir une problématique la recherche. Suite à ces différents constats, plusieurs méthodes de (géo-) visualisation ont été étudiées, développées avec notre plateforme FishEye, et présentées devant un comité de thèse en fin de première année.

Au cours de la deuxième année, nous avons approfondi la question de la modélisation de la visualisation et des méthodes d'analyse géovisuelle. Pour cela, nous avons mis au point un modèle ontologique qui puisse répondre aux nombreuses questions d'usage de l'information géographique et de la visualisation, appliqué au domaine maritime. Différentes règles ont ainsi été mises au point, à partir de ce modèle ontologique.

Au cours de la troisième année, nous avons développé un prototype d'utilisation de cette base de connaissances. Cet outil permet ainsi d'utiliser ce système sans connaitre la structure de l'ontologie, de manière interactive et intuitive. Ce module d'utilisation du système à base de connaissance a ensuite pu être intégré à la plateforme FishEye.

Structure de la thèse

En introduction, nous présentons le contexte d'application de cette recherche, qu'est l'analyse du trafic maritime, afin d'améliorer la sécurité et la sûreté de la maritimisation de l'énergie. Cette partie permet alors d'identifier les limites actuelles des systèmes de contrôle, aussi bien d'un point de vue opérationnel que d'un point de vue scientifique sur la représentation et l'analyse de l'information géographique maritime.

La première partie de cette thèse se consacre à l'état de l'art dans les domaines de gestions des risques et de la visualisation d'information géographique. Le chapitre 1 porte essentiellement sur les différentes méthodes, outils et organisation pour la gestion des risques. Le chapitre 2 propose alors un état de l'art relatif au processus d'exploration et d'analyse d'informations de mouvement, grâce à

l'analyse géovisuelle. Nous présentons notamment les travaux de formalisation de la connaissance dans le domaine de la visualisation d'information, afin de fonder notre recherche.

Dans le chapitre 3, nous présentons notre proposition d'un environnement d'aide à l'analyse géovisuelle, par l'utilisation d'un formalisme ontologique et de règles de raisonnement. Ce travail original dans les sciences de l'information géographique souligne notamment l'importance d'adopter une telle approche formelle pour décrire les méthodes de visualisation d'information géographique. Enfin, dans le chapitre 4, nous illustrons l'utilisation d'un tel système pour l'analyse de comportements à risques dans le domaine maritime. Nous comparons ainsi l'approche courante des risques maritime, du point de vue des contrôleurs ou des officiers de navigation, à celle proposée par notre approche formelle.

Enfin, dans la conclusion, nous revenons sur les apports de cette thèse aux domaines de la gestion des risques maritimes et de l'analyse géovisuelle. Nous présenterons alors plusieurs perspectives d'amélioration pour notre travail de recherche, ainsi que pour le domaine de l'analyse géovisuelle dans un contexte de contrôle des risques.

CHAPITRE 1 - L'ETRE HUMAIN FACE AUX COMPORTEMENTS A RISQUES

L'analyse des mouvements d'objets mobiles peut avoir de nombreuses fins d'utilisation, telles l'optimisation d'un espace géographique, ou la reconnaissance et la gestion des risques. Elle prend alors part à un processus plus long complexe, auquel elle vient apporter une contribution majeure pour la compréhension de la dynamique d'un système. Dans le premier chapitre de cette thèse, nous présentons de manière générale la notion de risques, ainsi que les différentes étapes dans la gestion et de l'analyse de ces risques, dans lesquels l'analyse de mouvements peut jouer un rôle.

La première partie de ce chapitre introduit la notion de risque et les différents modèles qui ont été proposés pour l'analyse, la gestion et la prise de décision en situation de crise. Les rôles complémentaires de la technologie et de l'être humain sont soulignés. La deuxième partie présente un état de l'art des études menées sur l'acceptation de nouvelles technologies, afin de comprendre les facteurs qui peuvent avoir une influence sur l'utilisation effective ou non d'outils, en temps de gestion de risques. Enfin, la troisième partie introduit le rôle de l'information géographique et de la cartographie dans le domaine de la gestion de crises, de risques, notamment pour l'étude de comportements au sein d'un système dynamique.

1.1. De la gestion des risques à l'aide à l'analyse

1.1.1. Les processus de gestion des risques

Nous retrouvons plusieurs définitions du risque, qui donnent généralement un sens dual à ce concept. Le risque peut être vu comme un concept dual, qui comporte à la fois une dimension **quantitative**, par les pertes potentielles (vulnérabilité), et une dimension **quantitative et probabiliste** avec la probabilité que survienne un évènement dangereux (Dufour et Pouillot 2002). Dans le cadre de la maîtrise des accidents majeurs relatifs à des substances dangereuses, la directive européenne du 9 décembre 1996 (Seveso II) propose une définition du risque.

> *Probabilité qu'un effet spécifique se produise dans une période donnée ou dans des circonstances déterminées. En conséquence, un risque se caractérise selon deux composantes : la probabilité d'occurrence d'un événement donné, la gravité des effets ou conséquences de l'événement supposé pouvoir se produire (Journal Officiel 1997)*

La notion de probabilité se retrouve dans la majeure partie des définitions du risque. Une manière plus graphique d'appréhender le risque est la courbe de Farmer, régulièrement utilisée en gestion de risques. La courbe de Farmer (voir Figure 1-1) a été mise au point par l'ingénieur nucléaire anglais Farmer : celle-ci permet de classer les risques selon leur occurrence et leur gravité (Farmer 1967). Les risques se répartissent selon une courbe décroissant en fonction de la gravité de ceux-ci, leur

fréquence suivant donc une loi inverse. Trois domaines sont alors identifiés, comme nous le montre ce schéma : les risques à fréquence très élevés mais à faibles conséquences ; les risques à fréquence moyenne mais de gravité importante ; enfin, les risques majeurs (évènements incertains dont les conséquences négatives sont considérables) qui doivent être absolument évités. Sous cette courbe se situe la zone de l'acceptable, mais au-dessus de la courbe, il s'agit de la zone des risques non souhaités, interdits, telles que peut l'identifier l'industrie du nucléaire (International Atomic Energy Agency 2014).

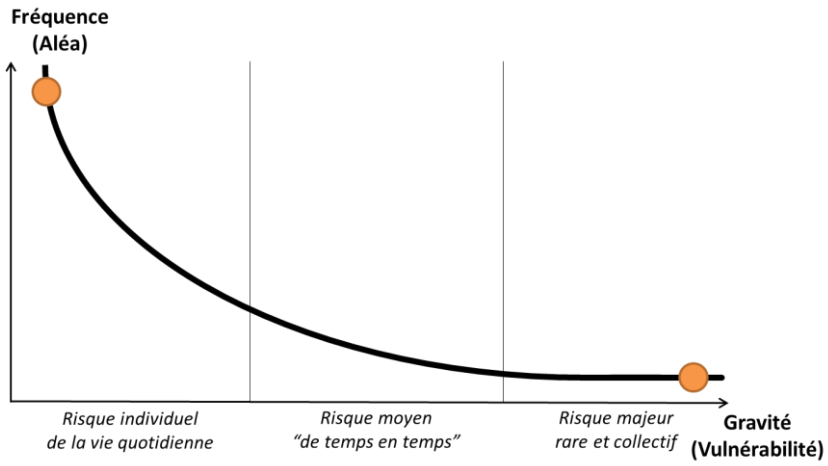


Figure 1-1. Courbe de Farmer : gravité et probabilité des risques

Les deux axes de ce graphique sont régulièrement repris pour définir les risques, ainsi que les politiques de prévention et de gestion des risques et crises. En 2005, le secrétaire à la Sécurité intérieure des États-Unis, Michael Chertoff, définit l'ensemble des risques comme une matrice à trois dimensions. Cette matrice prend en compte : les menaces potentielles, la vulnérabilité des installations ou des organisations, et les conséquences qui peuvent être subies (McNicholas 2011). Cette définition du risque vient prolonger la définition de l'ISO du guide 73 (ISO 2009), selon lequel un risque est la combinaison de la *probabilité d'un évènement et de ses conséquences* (référence directe à la courbe de Farmer). Cette idée de matrice à trois dimensions montre la haute complexité des informations à prendre en compte lors de la modélisation des risques. On retrouve cette formule assez régulièrement, à la fois dans le monde opérationnel et dans le monde académique (Moteff 2005). Toutefois, son utilisation est encore sujet à plusieurs débats, sur l'optimisation de la réduction des risques (Cox 2008).

Equation du risque : $Risque = Menace \times Vulnérabilité \times Impact$

En intégrant la réponse des contremesures qui viennent réduire les risques, l'équation du risque est alors compensée par cette nouvelle formule (Dubée 2001) :

$$Risque = \frac{Menace \times Vulnérabilité}{Contremesures} \times Impact$$

Cette équation montre le rôle important des contremesures, qui viennent réduire les conséquences directes des risques, sur un certain système. Dans la sous-partie suivante, nous présentons les différentes stratégies qui peuvent être mises en place pour la réduction des risques.

1.1.1.1. Différentes stratégies pour réduire le risque

La Figure 1-2 représente les différentes stratégies qui peuvent être mises en place afin de réduire les risques courus par un système. Les nombreux risques encourus sont organisés selon deux axes qui les qualifient : leur vulnérabilité (axe vertical) et leurs conséquences (axe horizontal). Ces deux axes reprennent deux des trois dimensions proposées par Chertoff, présentées plus haut. Les risques sont alors divisés en plusieurs groupes d'importance. Cette figure s'inspire de la courbe de Farmer, qui évalue l'acceptabilité des évènements (risques) selon leur fréquence de survenue et leur gravité.

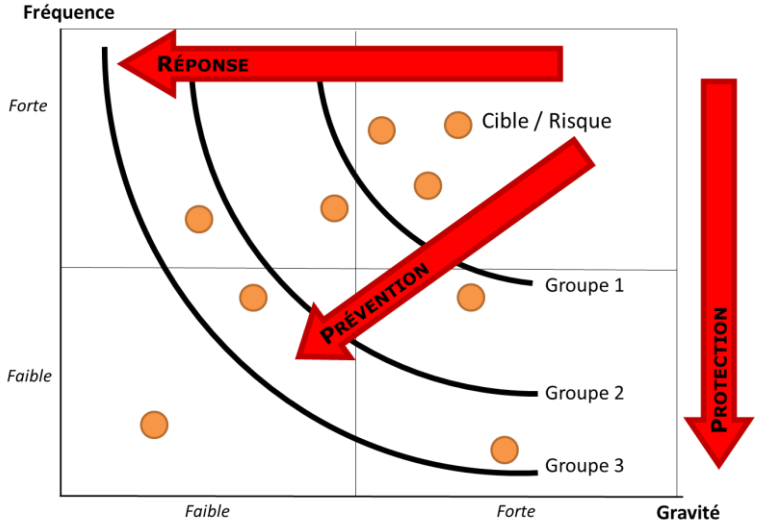


Figure 1-2. Différentes stratégies de réduction des risques (d'après *United States Coast Guard*)

Les éléments du groupe 1 sont les objets à plus hauts risques, qui doivent être considérés en priorité la plus haute lors de la prise de décision. Le groupe 2 correspond aux risques qui seront considérés plus tard, de manière moins urgente que le premier groupe. Le groupe 3 peut être, selon le cas, un risque à gérer ou bien à accepter. Enfin, le groupe 4 correspond aux risques acceptables, qui auraient été définis en amont. Une stratégie de réduction des risques représente alors l'ensemble des décisions qui feront passer un risque dans un groupe de moindre importance. En résumé, le groupe 1 est la zone de l'inacceptable ; le 2 et 3 la zone de crise ; le 3 la zone de l'acceptable.

Les différentes stratégies qui peuvent être mises en place face à un risque sont alors représentées par les trois flèches rouges du schéma : Réponse adaptée, Prévention des risques, Mesures de protection. Ces stratégies sont proposées par le modèle américain *Maritime Security Risk Analysis Model* (MSRAM), mis en place depuis 2005 par l'*U.S. Coast Guard* (service américain militaire pour le contrôle des frontières maritimes) (Adler et Fuller 2007).

Ces trois scénarios interviennent dans la gestion de risques, grâce à des outils d'aide à la décision. Ils permettent alors de gérer un certain risque, ou de le déplacer dans une catégorie de moindre importance. Répondre à un risque de manière adapté permet de réduire les conséquences, mais n'évitera pas une répétition de ce risque par la suite. Autrement, le scénario de prévention permet de réduire la vulnérabilité aux risques et l'attraction d'une certaine cible : cette décision, plus durable, permet de réduire la répétition d'un certain risque ainsi que ses conséquences. Enfin, le troisième type de gestion du risque est la mise en place de mesures de protection, telles que placer des gardes sur les navires. Cette solution permet de réduire fortement la vulnérabilité du système, et

donc sa probabilité. Cette option s'est vue être acceptée très récemment en France dans le domaine maritime, avec par exemple l'autorisation de gardes armés à bord des navires, pour prévenir les actes de piraterie (Le Monde 2013 ; Dumouchel 2013).

Le rôle de l'analyse des mouvements de navire, qui oriente notre recherche, n'a pas pour rôle de diminuer la gravité des accidents, mais plutôt de mieux comprendre le trafic afin d'éviter de tels évènements. Notre recherche s'inscrit donc à la fois dans une **stratégie de réponse adaptée** face au risque et de **prévention des risques** par l'étude du trafic maritime. Les mesures de protection peuvent être un effet de résultats d'analyse. Afin de gérer les réponses et les stratégies de prévention, le processus d'analyse n'est pas linéaire mais demeure au sein d'une boucle de contrôle. La sous-partie suivante présente des travaux effectués pour décrire un tel type de boucle de contrôle, au sein d'un processus de décision militaire.

1.1.1.2. La boucle OODA : les cycles de décision militaires

A des fins militaires, le stratège américain John R. Boyd invente dans les années 60 un cycle de contrôle pour gérer le cycle des décisions en combat aérien (Boyd 1996). Ce cycle de Boyd, appelé aussi la boucle OODA, est composé de quatre phases successives : Observation, Orientation, Décision et Action (*Observe, Orient, Decide, and Act*). La Figure 1-3 représente la boucle OODA et ses nombreuses composantes, telle qu'elle a été proposée à l'origine pour une utilisation militaire.

- La phase d'**observation** correspond à l'acquisition et fusion des données « terrain », relatives au système contrôlé et aux évènements qui s'y déroulent.
- La seconde phase, l'**orientation**, est la base de la prise de décision. Il s'agit de la mise en commun de nombreuses personnes et de connaissances, au sein d'un système socio-culturel complet. Ce système est alors constitué de traditions, d'expériences et d'habitudes variées. Ce système hétérogène est souvent complexe à gérer, car la prise de décision dépend de la mise en commun de nombreux points de vue.
- La phase de **décision** consiste à choisir une hypothèse réalisable, pour gérer le cours d'un évènement. Cette décision aura été sélectionnée parmi plusieurs autres possibilités, jugées alors moins bonnes pour le cas d'utilisation.
- Cette hypothèse est ensuite testée et exécutée lors de la phase d'**action**. Une fois l'action effectuée, des changements seront provoqués sur le système contrôlé, nécessitant alors un retour d'expérience. Les nouvelles observations sont prises en compte, ce qui implique un nouvelle orientation et de nouvelles décisions.

La Défense française souligne le rôle fondamental des interfaces homme-machine (IHM) afin de raccourcir la durée de la boucle OODA, afin de limiter l'effet des limites cognitives de la perception humaine (Observatoire du Monde Cybernétique 2013). Le monde militaire doit faire face aux très nombreuses données, qui se trouvent être utilisées dans la boucle OODA et qui doivent donc être visualisées par l'homme. Il est alors fondamental de stimuler la compréhension visuelle de l'information par des interfaces cognitivement adaptées.

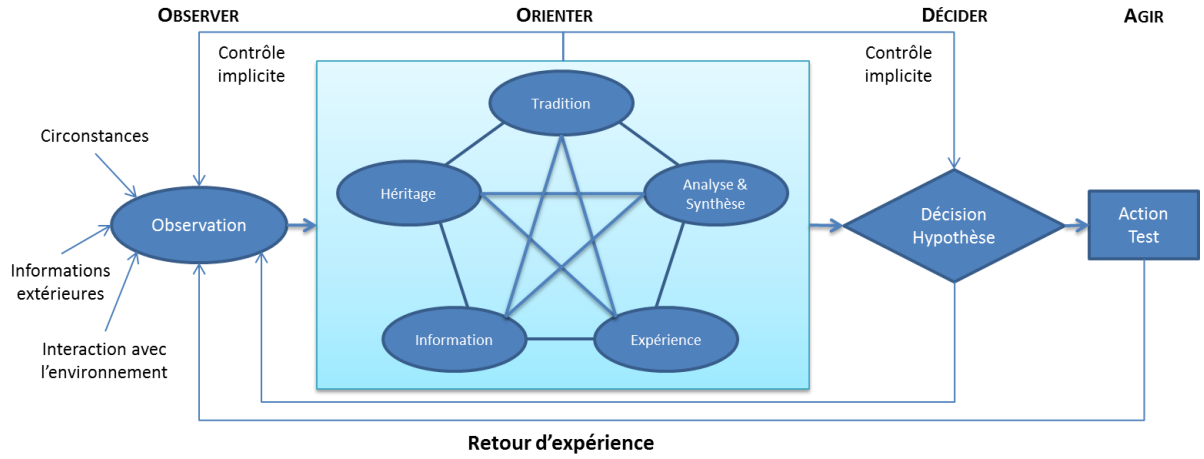


Figure 1-3. Boucle OODA, ou cycle de Boyd, pour la prise de décision dans le domaine militaire

La limite de ce modèle de décision se trouve néanmoins à la frontière entre Orienter et Décider. En effet, une équipe de décideurs présentant de nombreux profils différents peut bloquer l'adoption d'une décision commune. S'en suit alors de nombreuses itérations entre l'observation et l'orientation, sans pouvoir passer à la phase de la décision. Cette limite est aussi connue comme le problème OO-OO-OO (Ullman 2007).

Malgré cette limite humaine à l'utilisation de ce cycle, cette boucle pour la prise de décision a joué un rôle majeur dans la tenue de situations et la maîtrise des risques. La partie suivante présente des modèles proches de la boucle OODA, proposés pour la maîtrise des risques dans le sens général. Nous pouvons ainsi constater que l'héritage militaire guide le contrôle des situations les plus risquées, quel que soit leur contexte.

1.1.1.3. Anticiper les risques pour mieux les gérer

Lors de sa conception, la boucle OODA était alors destinée à une utilisation pour les décisions militaires, notamment par les pilotes d'avion. Toutefois, il est intéressant de constater que son utilisation peut s'appliquer à la gestion des risques en général. Avant d'introduire la gestion des risques, il est nécessaire de définir le concept de **contrôle** qui sera utilisé par la suite : celui-ci est au cœur des nombreuses recherches menées sur la gestion des risques. Nous pouvons donner la définition proposée par Hollnagel et Woods, qui définissent le contrôle de la façon suivante :

> *The ability to direct and manage the development of events and to compensate for disturbances and disruptions in a timely and effective manner. (Hollnagel et Woods 2005)*

En français, le contrôle est donc défini comme étant la capacité à diriger et gérer le cours des évènements, et à compenser les perturbations de manière rapide et efficace. Il s'agit donc de détecter, quantifier et interpréter les différences entre l'état observé du système d'intérêt et un état « idéal » désiré (modèle dit normal).

Le modèle OODA se reflète dans les travaux menés par Wybo, dans le domaine de la maîtrise des risques. Dans le cadre de l'organisation humaine en cas de crises, Wybo (Wybo 2012) propose une boucle de quatre étapes successives pour décrire la gestion des risques. Ces étapes sont illustrées sur la Figure 1-4. Ce processus en quatre phases permet la découverte de nouvelles connaissances des

risques, qui sont ensuite utilisées pour améliorer la gestion des évènements. Ces quatre phases successives sont :

- ***Vigilance :*** Cette étape demande un contrôle du système par différents moyens de surveillance (indicateurs, cartes, etc.). Il s'agit généralement de données quasi-temps réel qui représentent l'évolution d'une situation, d'objets mobiles (le trafic maritime, par exemple).
- ***Gestion de l'urgence :*** Débute lorsqu'un évènement arrive, sans avoir pu être prédit. De nouvelles solutions doivent être mises en place afin de stabiliser le système. C'est ici qu'intervient la notion de contrôle, telle que définie par Hollnagel et Woods (Hollnagel et Woods 2005) : la capacité à diriger et gérer le cours des évènements et à compenser les perturbations rapidement et efficacement.
- ***Apprentissage :*** Les retours d'expérience sont collectés et analysés, afin d'être utilisés par les phases suivantes. Ils permettront une meilleure gestion des risques à venir, si ceux-ci sont similaires aux évènements passés. Ces risques peuvent alors mieux être anticipés.
- ***Anticipation :*** Résultat des retours d'expérience et de l'analyse des évènements passés, de nouvelles connaissances sont créées et prises en compte pour améliorer le contrôle sur le système d'intérêt. Il s'agit de construire des connaissances à partir des expériences passées, afin de mieux comprendre et gérer les situations à venir.

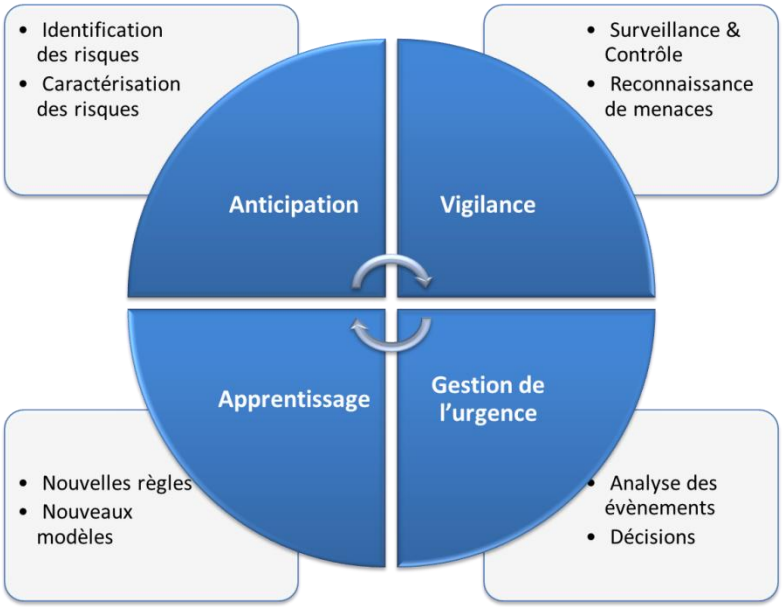


Figure 1-4. Processus de maîtrise des risques (d'après Wybo 2012)

Cette boucle de maîtrise des risques, qui peut s'appliquer dans tout type de système social, socio-technique, etc., s'appuie sur l'importance de l'anticipation qui se base sur les évènements passés. Cette phase d'anticipation est le cœur de nombreux projets actuels pour la gestion des risques : par exemple, les industriels (énergie nucléaire, énergie électrique), des risques organisationnels (risques humains, défaillance) ou des risques de transport (maritime, aérien, routier). Elle s'appuie sur une compréhension des risques courus par le système d'intérêt et leur modélisation basée sur les observations et retours d'expérience (similaire à la phase Observation dans OODA).

Dans le cas de ce modèle, la place de l'apprentissage et de l'anticipation montre la constante évolution dans laquelle s'inscrit ce processus de maîtrise des risques. Il n'existe pas un catalogue complet des risques encourus et des solutions convenues : tout comme dans la boucle OODA, la gestion des risques est un cycle constant qui se complète lui-même, basé sur les observations du système, des évènements et des résultats des actions entreprises. L'anticipation est une phase de modélisation des connaissances du domaine d'application, qui permet de poser les risques connus, susceptibles de se reproduire, et la façon de les gérer.

D'une manière similaire à Wybo, Hollnagel identifie quatre étapes majeures qui doivent être assurée lors du contrôle d'un système socio-technique donné (Hollnagel 2004, 185). Il s'agit là aussi de poser les connaissances issues d'un certain contexte, de manière à prévoir les risques encourus et les réponses adaptées. Ces quatre étapes sont, dans l'ordre :

- ***Caractériser le contexte :*** milieu de travail, but, ressources, etc.
- ***Identifier les risques :*** scénarios possibles, tâches, perturbations possibles, etc.
- ***Analyser les risques :*** vulnérabilité / probabilité, conséquences, etc.
- ***Décider des contre-mesures :*** défense, surveillance, procédures, communication, etc.

Dans chacun de ces modèles pour la gestion des risques, nous pouvons voir que ceux-ci sont rarement prédits exactement : seule une anticipation des nombreuses possibilités et de leur occurrence est à prendre en compte, de manière à anticiper les menaces. Il est aussi nécessaire de savoir comment **reconnaitre** ces évènements, de manière à prendre les décisions adaptées.

Le modèle décrit par Wybo décrit avant tout l'organisation humaine dans un contexte de maîtrise des risques. Sur ce modèle, on peut retrouver le **rôle des contrôleurs** maritimes, qui s'intègre à l'étape de vigilance et de gestion de l'urgence, grâce au support des **décideurs**. Aussi, les deux autres étapes viennent appuyer le **rôle joué par les analystes**, pour l'étude d'un évènement passé, et des chercheurs qui participent à la modélisation des risques. Cette approche montre donc son intérêt dans la séparation des rôles des acteurs de la sécurité et sûreté maritime, tels que nous les avons présentés dans la partie *Les acteurs de la « Sécurité et Sûreté de la Maritimisation de l'Energie »* en introduction de ce manuscrit.

Toutefois, de façon à pouvoir gérer des risques nombreux et variés de manière efficace, l'homme nécessite des outils avec lesquels mener les observations, l'analyse, la modélisation des risques et la prise de décision. Il est fondamental que ces outils dont l'homme est usager soient adaptés aux limites cognitives de l'homme, car l'erreur humaine dans l'utilisation de tels outils de contrôle est toujours une cause de risques très importante (Noyes 2001). Dans la partie suivante, nous présentons des travaux de recherche effectués dans l'ingénierie des systèmes cognitifs, notamment pour les systèmes de contrôle.

1.1.1.4. Les modèles cognitifs pour les systèmes de contrôle

La surcharge d'information pour la prise de décision par un être humain est un problème identifié depuis plusieurs décennies, notamment avec la métaphore de l'autoroute de l'information (Hiltz et Turoff 1985 ; Maes 1994). Il est naturel de constater qu'aujourd'hui même, au cours de ce que l'on appelle l'*Âge de l'Information*, les flux d'information disponibles augmentent de jour en jour, notamment grâce aux prouesses technologiques de ces dernières décennies dans les domaines de l'acquisition, le stockage et la diffusion de données (Hilbert et López 2011). Cette autoroute de

l'information nous mène au dilemme d'acquérir l'information beaucoup plus rapidement que nos capacités à pouvoir l'analyser (Keim et al. 2006).

Face à l'acquisition de ces grandes quantités de données et la limite de leur traitement et compréhension, la notion d'***ingénierie des systèmes cognitifs*** est proposée au début des années 80, notamment pour la conception des systèmes de contrôle (Hollnagel et Woods 1983 ; Noyes et Bransby 2001 ; Strater et al. 2006). Hollnagel et Woods soulignent alors trois principes fondamentaux qui caractérisent un système cognitif (Hollnagel et Woods 1983) :

- *Être focalisé sur un but défini et basé sur la manipulation de symboles*. Selon Harnard, la manipulation de symboles est la manipulation de formes symboliques, sans rapport à leur interprétation (Harnard 1994).
- *Être adaptatif*, proposer plusieurs points de vue pour un même problème.
- *Se baser sur la connaissance de l'utilisateur et de l'environnement*, de manière à adapter les actions selon ce contexte.

A cette époque, la notion de formalisation des connaissances par un langage commun est alors déjà initiée, de façon à pouvoir proposer des systèmes cognitifs qui soient exploitables de manière opérationnelle. Ce vocabulaire commun permet essentiellement de soutenir le troisième point identifié par Hollnagel et Woods, relatif à la connaissance de l'utilisateur et de son environnement. Nous reviendrons sur cette étape de formalisation de la connaissance au cours du chapitre 2 de ce manuscrit.

Les systèmes cognitifs, tels que présentés ici, doivent donc permettre le contrôle d'une situation grâce à la collaboration entre l'homme (opérateurs, analystes, décideurs) et la machine (calculs, visualisations et interactions). Les connaissances et l'expérience de l'homme doivent être soutenues par la puissance de calcul des machines, de manière à alléger la charge cognitive, due aux nombreuses informations et tâches de travail. Suite aux travaux sur l'ingénierie des systèmes cognitifs menés par Hollnagel et Woods dans les années 80 (Hollnagel et Woods 1983), c'est à la fin de cette même décennie que Woods, Roth et Bennet introduisent la notion de ***système cognitif joint*** (*Joint Cognitive Systems* ou JCS) (Woods et al. 1990). Les JCS sont alors introduits de manière à décrire les processus cognitifs, lorsqu'un être humain interagit avec une représentation de l'information à étudier (par exemple, un outil de visualisation d'information).

Ces systèmes cognitifs joints permettent de mutualiser les « ressources humaines » et les « ressources machines », sur le plan fonctionnel. Ces deux ressources, posent alors les bases d'un ***système socio-technique*** (Chatzis 1993 ; Kéfi et al. 2004). Celles-ci doivent interagir lors de l'exploration d'information, de l'analyse d'information et de la prise de décision. Rappelons une définition du système socio-technique, illustré sur la Figure 1-5.

> *Système composé d'éléments techniques (physiques) et organisationnels, formellement organisés sur la base de règles et de rôles, de normes et de contrôles, de programmes et de positions plus ou moins standardisées, en vue de réaliser des objectifs prédéfinis. (Chatzis 1993)*

Dans notre cas d'étude, un tel système socio-technique représente l'ensemble des acteurs de la sécurité et sûreté maritime, tels que nous les avons présentés en introduction, ainsi que le panel

d'outils qu'ils sont amenés à utiliser chaque jour pour le contrôle et l'analyse de cet environnement. Toutefois, les technologies utilisées actuellement ne peuvent pas être qualifiées de JCS, suite aux limites d'utilisation identifiées au début de ce manuscrit, introduisant une marge encore importante dans les rôles de ces deux aspects du système maritime.

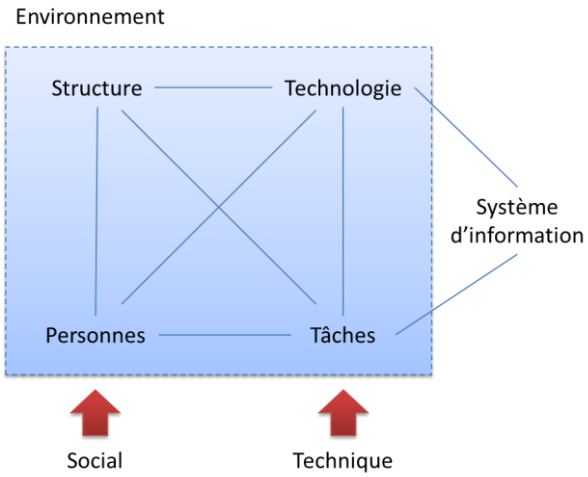


Figure 1-5. Modèlisation d'un système socio-technique et impacts des systèmes d'information sur ce système (d'après Kéfi et al. 2004)

Afin de décrire l'utilisation des JCS, Hollnagel propose en 1993 le modèle COCOM (*Contextual Control Model*) pour une description basique du contrôle d'un système (Hollnagel 1993). Ce modèle est basé sur les évènements qui sont successivement enregistrés depuis le système d'intérêt. Les actions à entreprendre dépendent du contexte, défini par ces évènements, contrairement à un protocole procédural qui définit le plan d'actions par une séquence prédéfinie de contremesures. Le modèle COCOM est illustré sur la Figure 1-6, mettant en évidence les trois étapes identifiées par les auteurs: les évènements enregistrés dans le système contrôlé, la mise à jour des concepts (par exemple, les modèles de comportements à risques), les actions à mener et identifiées par les décideurs.

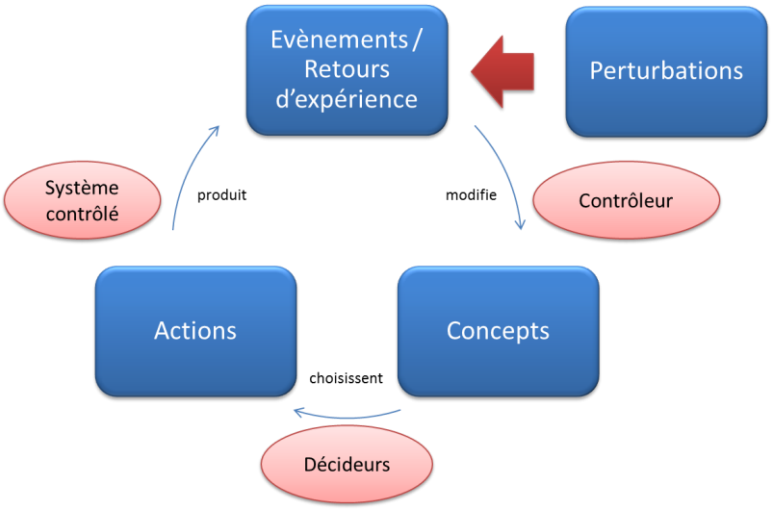


Figure 1-6. Modèle COCOM pour le contrôle des risques (d'après Hollnagel 1993)

Tout comme dans le cycle proposé par Wybo, nous pouvons retrouver sur le modèle ECOM le rôle des contrôleurs et des analystes, ici regroupés en une seule étape entre les évènements surveillés et les concepts utilisés. Le travail complémentaire de ces rôles, dans le domaine maritime, permet de

maintenir une connaissance importante des risques maritimes et des comportements qui correspondent à ces situations.

Toutefois, Hollnagel et Woods reprennent rapidement ce modèle du contrôle des risques, car COCOM simplifie trop le processus de contrôle par rapport aux situations réelles qui peuvent être observées sur le terrain (Hollnagel et Woods 2005 ; Hollnagel 2000). En effet, ces différentes étapes du contrôle de la situation peuvent avoir lieu de façon simultanée et en parallèle, tout en produisant de nouvelles informations pour les étapes suivantes. Ainsi, les mêmes auteurs proposent plus tard une nouvelle version améliorée et approfondie grâce au modèle ECOM (*Extended Control Model*) (Hollnagel et Woods 2005), afin de décrire les différents niveaux de contrôle. Contrairement au modèle précédent, les étapes de ce modèle ne sont pas exécutées séquentiellement mais sur des couches de contrôle parallèles et simultanées, au sein des JCS (voir Figure 1-7). Chacune de ces couches produit des résultats et des observations, qui sont ensuite utilisés dans les couches de plus bas niveaux. Dans cette superposition de couches de contrôle, le couplage de l'homme et de la machine opère à chaque niveau, grâce à des échanges d'informations entre les deux.

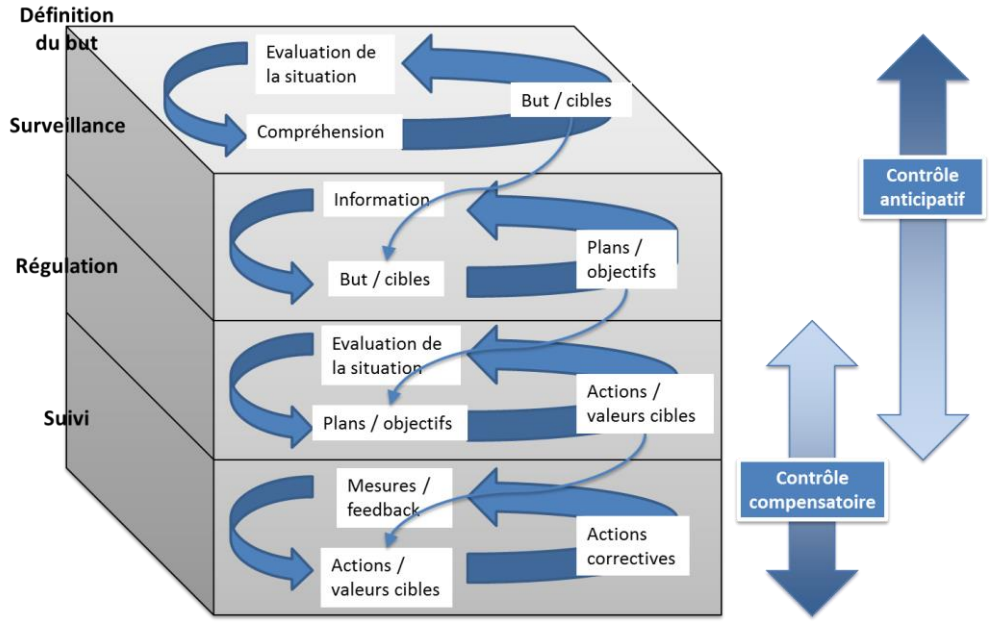


Figure 1-7. Modèle ECOM pour le contrôle des risques (d'après Hollnagel et Woods 2005)

L'exécution de chacune de ces couches de contrôle dépend donc des boucles supérieures, mais celles-ci ne demandent pas toutes le même type d'attention, la même fréquence et le même temps d'utilisation. Chaque couche de contrôle influe sur les couches de plus bas niveaux : l'identification des risques dépend du contexte, l'analyse effectuée dépend des risques possibles, etc. Ces quatre couches de contrôle sont, de la plus haute à la plus basse :

- ***Définition du but (Targeting) :*** Il s'agit de la phase préliminaire pour situer l'action de contrôle. On décide des cibles d'intérêt dans le système surveillé, les comportements que l'on attend pour mener à bien l'analyse du système, normaux ou anormaux.
- ***Surveillance (Monitoring)*** *:* A partir des informations reçues des éléments cibles, il s'agit ici de dresser une situation dite « normale » pour les cibles identifiées, et les observations actuelles effectuées sur le système. Cette situation normale est l'objectif du contrôle.

- ***Régulation (Regulating)*** *:* Application de la définition du contrôle, telle que donnée plus haut. Pour cela, l'état réel du système est comparé aux objectifs définis dans les niveaux supérieurs du contrôle. Selon l'évaluation de la situation courante, les objectifs identifiés doivent être appliqués de manière adaptés.
- ***Suivi (Tracking)*** *:* Afin de maîtriser les situations rencontrées au cours de la phase de régulation du système, des actions correctives sont menées pour recentrer les valeurs cibles qui sont régulées. Ces choix dépendent des mesures et retours du système (feedbacks).

Le modèle ECOM met en évidence deux types de contrôle, que sont le contrôle anticipatif et le contrôle compensatoire. Les deux couches d'actions du bas, la régulation et le suivi, correspondent à des actions et réactions de courte durée ; celles-ci font appel à l'expérience et aux compétences de l'homme en charge du contrôle. Ce contrôle est dit ***compensatoire***, car il s'agit de réagir rapidement à des situations réelles par des interventions directes. Les deux couches supérieures, la définition du but et la surveillance, correspondent à des actions de durées plus longues. Notamment puisque la définition du but a lieu dès la préparation du travail, puis pendant tout le processus de contrôle (Hollnagel et Woods 2005, 152). Pour cette raison, ce contrôle est dit ***anticipatif***.

Toutefois, le modèle ECOM ne montre pas l'influence des processus psychologiques du traitement de l'information, tels qu'ils ont pu être décrits dans le modèle SRK (*Skills, Rules, Knowledge*) de Rasmussen (Rasmussen 1983 ; Engström et Hollnagel 2007). Le modèle SRK organise ces processus de traitement selon la performance humaine (Rasmussen 1983), en distinguant les comportements basés sur les ***compétences*** (ou habileté), les ***règles*** ou les ***connaissances*** de l'opérateur. Face à une certaine situation, les compétences permettent de réagir de façon automatisée sans effort cognitif. Le comportement doit être basé sur des règles lorsqu'une situation rencontrée est plus complexe mais familière, et nécessite des procédures et règles définies. Enfin, la réaction d'un opérateur doit se baser sur ses connaissances lorsqu'un évènement inattendu est rencontré : pour cela, une analyse des faits et des informations doit être menée.

1.1.1.5. Bilan

Parmi ces nombreuses propositions de méthodes pour identifier et gérer le risque, on reconnait une trame commune à ces travaux. De ces propositions, il est important que nous identifiions la procédure de gestion dans un contexte technologique, c'est-à-dire la manière dont des outils d'analyse permettent de mieux gérer l'information, pour la compréhension des risques dans les données de mouvement. La Table 1-1 permet de synthétiser les différentes étapes proposées par ces travaux de la communauté du risque, afin d'en identifier les grands axes communs. A chacune de ces étapes, nous avons sélectionnées (✓) celles où s'intègre notre démarche grâce à l'analyse géovisuelle de comportements, définie par la suite, d'après la littérature sur le sujet. De cette manière, nous situons la place de notre recherche au sein de ces processus généraux de gestion des risques.

Afin de représenter ces étapes, qui ne sont pas nécessairement dans l'ordre du modèle original, nous avons choisi de se baser sur les axes proposés dans le modèle ECOM, que l'on retrouve dans les différentes stratégies de réduction des risques (partie 1.1.1.1 de ce chapitre). On peut en effet graduer ces termes de gestion de risques de l'anticipation (étude des données afin de mieux gérer le futur) à la compensation (action présente sur le terrain). On remarque que dans chacun de ces modèles de gestion des risques, il est fait mention d'une étape d'apprentissage des données passées

(accidents et retours d'expérience) afin d'améliorer les décisions futures, les modèles de comportements, la définition des risques eux-mêmes, etc. Cette étape de modélisation est alors fondamentale pour améliorer à la fois la reconnaissance future de ce type d'évènements, mais aussi les décisions à prendre en conséquence. Par cette étape de modélisation, on identifie un acteur du type « chercheur » qui vient alors s'intégrer à la boucle de contrôle, sans être nécessairement dans le processus de prise de décision ou de contrôle en temps réel du système. Il est important de relever ce rôle, qui nous sera utile dans la suite de cette recherche (rôle du chercheur pour la modélisation des risques et de l'information géographique).

En plus de cet acteur, on retrouve donc les rôles des personnes du domaine relatif aux risques gérés, tels que les contrôleurs et les décideurs, dans la phase d'action et de compensation en gestion des risques. Bien que le modèle OODA n'ait pas été proposé dans un contexte de gestion de risques, nous pouvons voir que ce modèle historique et militaire a influencé la culture de la gestion des risques sur les aspects théoriques et organisationnels.

Table 1-1. Bilan des étapes de la gestion des risques par la technologie

Etape / Modèle	Anticipation	.....		Compensation
OODA (années 60, Boyd 1996)	Observer ✓	Orienter ✓	Décider	Agir
COCOM (Hollnagel 1993)	Evènements système ✓	Mise à jour concepts ✓		Action
Contrôle système socio-technique (Hollnagel 2004)		Caractérisation du contexte ✓ Identification des risques ✓ Analyser les risques ✓ Identification des contremesures		
ECOM (Hollnagel et Woods 2005)	Surveillance ✓	Définition but	Régulation	Suivi

Dans cette partie, nous avons mis en évidence la complexité et les nombreuses stratégies pour la gestion des risques, face à une situation connue ou nouvelle. Dans les différents travaux présentés, le rôle de la « machine » pour soutenir le processus de contrôle par l'homme a toujours été mis en avant. Cela est la conséquence directe de la surcharge cognitive induite par la masse d'informations à traiter, la complexité des scénarios possibles et des décisions à prendre. Les JCS ont un rôle fondamental dans la collaboration homme-machine, pour une compréhension rapide et une prise de décision efficace en situation de contrôle. Dans la partie suivante, nous présentons des travaux qui ont été réalisés afin de faire cohabiter les connaissances de l'homme et la puissance de la machine pour l'analyse de situation et la prise de décision.

1.1.2. De l'aide à l'analyse à la prise de décision

Le concept de ***Situation Awareness*** (SA) provient de la psychologie aéronautique (Mogford 1997). Il correspond à la perception des éléments de contexte temporel et spatial, leur compréhension, ainsi que leur projection lorsqu'un élément est amené à changer (Tretmans et van de Laar 2013). Par la suite, nous gardons la notation SA pour renvoyer à ce terme, qui n'a pas d'expression équivalente en français. La SA a ensuite été très rapidement adoptée par les communautés du contrôle du trafic aérien (Foyle et al. 2005), du contrôle du trafic maritime (Riveiro et al. 2008), ou encore pour le *Command & Control* militaire pour les actions sur terre (Strater et al. 2006). La SA permet donc de donner de la cohérence à un ensemble d'évènements, de manière à prendre des décisions efficaces par la suite (Endsley 1995 ; Glandrup 2013). Ce concept a donc la triple vocation de permettre : la **perception** des données de l'environnement surveillé, la **compréhension** de l'information actuelle et l'**anticipation** de l'état futur du système (Chalandon 2003 ; Endsley 1995 ; Endsley et al. 2003).

Dans cette définition de la SA, on retrouve deux étapes fondamentales de la gestion des risques, dont nous avons vu les différentes propositions dans la partie précédente. La SA a donc un rôle majeur dans la gestion des risques, grâce à l'anticipation de la situation qu'elle induit.

Dans le domaine de la surveillance maritime, le rôle majeur de la SA est notamment souligné par Glandrup (Glandrup 2013), lorsqu'il cite les grandes limites des systèmes de surveillance actuels : le trop-plein d'information et le manque d'assistance dans leur analyse. Ces limites se retrouvent, par ailleurs, dans de nombreux autres types de systèmes de surveillance. Comme il l'écrit correctement, le grand danger est de ne pas identifier des évènements importants dans cette quantité de données.

> *The monitor and control systems that these organizations currently use cannot cope with the amount of information that is available. At this moment, the majority of information sources must be interpreted by operators, or simply ignored. (…) Today's maritime safety and security systems do not provide sophisticated interaction mechanisms to deal with the diversity of information. (Glandrup 2013, 23)*

Un système qui supporte la SA doit donc aider dans la détection et l'identification sur des situations anormales, et peut-être dangereuses, ainsi que permettre de reconnaitre de situations dites normales (Tretmans et van de Laar 2013). Boy qualifie alors la SA comme un indicateur de sécurité dans des systèmes hommes-machines hautement automatisés (Boy 2011, 12). Lorsque l'on parle de *Situation Awareness*, il s'agit donc d'un concept large qui représente les facteurs cognitifs dans l'aide à la compréhension d'une certaine situation. La SA soutient ainsi l'analyse et la prise de décision en améliorant la compréhension d'un système d'intérêt.

Les outils qui permettent de maintenir la SA présentent généralement des fonctionnalités de production, de partage et d'analyse des informations nécessaires à la prise de décisions. Ces informations sont de l'ordre de données sémantiques et spatio-temporelles, recueillies en temps réel pour informer au mieux de l'état de l'environnement surveillé (Goodchild 2010 ; Seppänen et Virrantaus 2010).

Une fois présenté le concept de *Situation Awareness* dans sa globalité, nous pouvons facilement comprendre le concept reconnu de ***Maritime Domain Awareness*** (MDA). L'Organisation Maritime Internationale a défini le MDA comme étant la compréhension de toutes les informations et activités

associées au milieu maritime, qui puisse avoir un impact sur la sécurité, la sûreté, l'économie ou l'environnement (IMO 2013). Il s'agit donc d'une application de la SA au domaine maritime. Dans les sous-parties qui suivent, nous présentons des modèles de systèmes pour l'aide à l'analyse, ainsi que les technologies qui peuvent être employées pour mettre au point de tels environnements.

1.1.2.1. Les systèmes d'information pour l'aide à l'analyse

Parmi les systèmes qui servent à l'aide à l'analyse d'information, deux grands types sont récurrents dans la littérature. Les systèmes d'aide à la résolution de problèmes et les systèmes d'aide à la décision. Ceux-ci peuvent être vus comme deux concepts d'outils complémentaires dans l'analyse de grande quantité d'information. Nous présentons ici ces deux concepts et leurs apports, quelques domaines d'applications, ainsi que leurs limites.

1.1.2.1.1. Les environnements d'aide à la résolution de problèmes

Le concept d'***environnement (ou système) d'aide à la résolution de problèmes***, ou bien avec son équivalent anglais qui est plus souvent lu *Problem Solving Environment* (PSE), remonte au début des années 1990. Ce sont des environnements informatiques, logiciels, se basant sur des méthodes de calcul automatisées dans des interfaces homme-machine, afin de guider dans la résolution de certains problèmes (Gallopoulos et al. 1992). Ils se basent généralement sur des outils de simulation pour prédire l'évolution d'un système (Gallopoulos et al. 1994).

Selon Gallopoulous, un PSE est utilisé pour sélectionner les meilleures solutions à un problème donné, tout en conservant un historique des tâches menées par l'utilisateur (Gallopoulos et al. 1992). Dans ce même article, les auteurs identifient deux raisons qui ont motivé l'utilisation de PSE : (1) **résoudre des problèmes de manière rapide** et (2) **mener des actions qui n'étaient alors pas identifiées précédemment**. Pour cela, des équipes inter-disciplinaires doivent être mises en place de manière à produire ce type de système expert, notamment grâce à un travail de formalisation de la connaissance des experts vers un système intelligent (Gallopoulos et al. 1994).

Des exemples de PSE peuvent se trouver dans le domaine des mathématiques et statistiques (par exemple, SAS, R, Matab), de l'architecture et l'emménagement (logiciels de CAD), ou de la géographie (logiciels SIG). Les travaux par Jung et al. illustrent la création d'un PSE moderne pour la résolution de problèmes liés à des données géographiques, basé sur une formalisation des connaissances par ontologies (Jung et al. 2013). Ce modèle permet de résoudre des problèmes topologiques par l'utilisation de service web pour l'analyse de données spatiales, selon leur projection, leur géométrie et les fonctionnalités proposées par les SIG. En retourne alors une succession d'opérations de traitement spatial pour l'analyse des données (buffer, recherche d'intersection, etc.). Toutefois, dans la définition originale des PSE, le contexte d'utilisation et le profil de l'utilisateur ne sont pas pris en compte : le problème va être résolu par rapport aux données qui sont disponibles, ainsi que les différentes étapes observées dans la résolution habituelle de ce problème. Par exemple, le PSE géographique proposé par Jung et al. se base sur les fonctionnalités des SIG, et propose la meilleure stratégie d'utilisation d'un tel logiciel pour des questions identifiées. Les compétences et limites de l'utilisateur ne sont pas prises en compte dans cette proposition.

1.1.2.1.2. Les systèmes d'aide à la décision

Les ***systèmes d'aide à la décision (SAD)***, ou bien *Decision Support Systems* (DSS), sont des systèmes d'information qui supportent l'analyse et la prise de décision dans des situations complexes. Pour

cela, tout comme les PSE, ils combinent à la fois les connaissances de l'homme et le processus d'analyse, ainsi que la puissance de calcul des machines.

Les SAD permettent de retrouver et d'accéder les données nécessaires à la décision, et proposent des techniques d'analyse grâce à des interfaces interactives et simples d'utilisation (Sprague 1980). Ces SAD sont souvent fondés sur des connaissances formalisées, qui proposent une expertise par rapport à une classe de problèmes à partir de faits et règles (Power 2002). De plus, ils ont la spécificité d'être **flexibles** et **adaptables** selon le contexte d'utilisation, selon Sprague (1980). Cette flexibilité permet en effet de s'accommoder aux décisions menées par l'utilisateur, ainsi qu'à son environnement de travail.

Un SAD se base sur plusieurs composantes pour la formalisation et l'utilisation des connaissances : un modèle de l'environnement ciblé, un moteur de simulation, un modèle d'utilisateur et un moteur de recommandation qui sert d'interface entre l'homme et la machine. Il s'agit donc de systèmes experts pour guider dans la prise de décision, grâce à l'étude d'un savoir formalisé. Selon la définition du rôle des systèmes experts, il est donc aussi nécessaire qu'un SAD puisse fournir des justifications aux choix et décisions qu'il effectue (Jackson 1999).

De nombreux exemples de SAD dans le domaine de la décision à partir de données spatiales se retrouve dans la littérature (Andrienko et al. 2007 ; Pala et Wilson 2013). Ces systèmes d'aide à la décision viennent supporter l'utilisateur dans l'étude d'un territoire géographique, avec la visualisation de nombreux indicateurs par rapport à un système sensible : crues et inondations, diffusion de produits nocifs, etc. L'étude des données spatiales va jouer un rôle fondamental dans la prise de décision et le choix des plans, scénarios, interventions, etc. Le SAD permet aussi de faire évoluer en temps réel et de manière intelligente les interventions des agents sur le territoire, par exemple de la gestion de produits dangereux (Andrienko et Andrienko 2007).

Dans le domaine spatial, les SADARS (systèmes d'aide à la décision à référence spatiale) sont généralement le couplage d'un SAD à un SIG, pour la représentation des données spatiales (White et Engelen 2000 ; Andrienko et al. 2007 ; Mendas et al. 2007 ; Densham 1991). On retrouve notamment dans de tel système d'aide à la décision spatiale des outils de simulation, permettant d'anticiper l'état d'un objet, du système spatial étudié, ou de son impact sur l'environnement concerné, suite à un évènement sensible ou une crise majeure (Fournier 2005 ; Claramunt et al. 2007 ; Ahola et al. 2007 ; Florez Guzman 2012).

Un autre domaine d'intérêt pour les SAD est de l'ordre militaire : l'aide à la décision au sein de données spatiales et sémantiques peut aider à l'évaluation de menaces ennemies, grâce à des outils de simulation et d'analyse proposés par de tels environnements (Dahlbom et Helldin 2013). Pour cela, les SAD doivent utiliser de nombreuses sources d'informations et les présenter de manière claire et efficace à l'utilisateur, afin de permettre la meilleure analyse possible.

1.1.2.2. Les technologies pour supporter l'analyse

De manière à mettre en place des systèmes d'aide à la décision, ou d'aide à l'analyse, de nombreuses technologies ont été identifiées et ont montré leur efficacité jusqu'aujourd'hui. Ces technologies peuvent être utilisées dans le cadre d'un système dit *intelligent*, c'est-à-dire calquant le processus de raisonnement humain, ou bien servir de support à l'analyse guidée par l'utilisateur. Dans cette partie, nous présentons quelques technologies adoptées pour l'aide à la décision ou l'aide à l'analyse, ainsi

que certains travaux correspondant. Ces nombreuses technologies sont souvent employées de manière complémentaires dans ces systèmes.

1.1.2.2.1. Réseaux bayésiens

Les réseaux bayésiens sont utilisés dans le cas où des probabilités doivent guider la prise de décision, suite à des incertitudes. Cette théorie est basée sur le théorème de Bayes, qui permet de calculer la probabilité de A sachant B, si l'on connait la probabilité de A, de B et la probabilité de B sachant A. Un réseau bayésien est formé de nœuds, qui représentent des facteurs, évènements ou décisions. Ces nœuds sont liés au sein d'un graphe orienté et pondéré selon les probabilités. La Figure 1-8 donne un exemple simple de réseau bayésien appliqué à la lutte contre la piraterie, et composé de trois nœuds : selon les différents paramètres tels que la saison et la présence d'une attaque de pirates, les dégâts causés sur la plateforme offshore seront plus ou moins importants. Le réseau bayésien permet ainsi de donner la probabilité que la probabilité de réussite d'un nœud final (présence de dégâts), selon les paramètres observés. Ce travail a été étudié en détail par Bouejla et al. pour le domaine de la piraterie maritime (Bouejla et al. 2014).

Attaque ...	Saison O...	Oui	Non
Oui	Oui	100,000	0,000
	Non	90,000	10,000
Non	Oui	100,000	0,000
	Non	5,000	95,000

Figure 1-8. Exemple de réseau bayésien à trois nœuds, ainsi que les probabilités correspondantes

Un réseau bayésien peut donc être utilisé pour modéliser un processus de prise de décision, dans un cas probabiliste et non pas déterministe (règles si, alors). De nombreuses applications sont retrouvées dans le domaine de la fouille de données (Liebig et al. 2009) ou l'analyse et la gestion de risques (Neerincx et al. 2009 ; Bouejla et al. 2012 ; Chaze et al. 2013). L'exemple du projet SARGOS pour une réponse graduée face à une attaque de pirates prouve l'efficacité d'une telle technologie pour une prise de décision adaptée (Chaze et al. 2012 ; Chaze et al. 2013).

Les nombreux travaux sur les réseaux bayésiens mettent en avant l'importance de la modélisation des informations et des processus associés à la prise de décision, notamment dans un cadre probabiliste. Si l'on considère une approche déterministe pour l'analyse de l'information et la prise de décision, les nouvelles technologies de moteurs de règles peuvent jouer un rôle tout aussi important que les réseaux bayésiens.

Un tel type de réseau s'appuie donc sur des probabilités, issues de faits, d'observations et d'incertitudes. Le réseau ainsi constitué est donc un arbre de décision qui permet un raisonnement automatisé par le système. Dans la prochaine sous-partie, nous présentons un autre mode de formalisation et de raisonnement, mais qui ne s'appuie pas sur des probabilités. Au contraire, le formalisme ontologique va permettre de mettre en place des règles à prédicats unaires et binaires, sans probabilités pour pondérer ces liens.

1.1.2.2.2. Modélisations ontologiques et moteurs de règles

Les ontologies informatiques ont été introduites à la fin des années 90 et définies comme une *spécification explicite d'une conceptualisation partagée* (Gruber 1993 ; Gruber 2000). Une ontologie est un modèle composé de concepts (ou classes), de relations qui les lient (inclusions *est-un* ou sémantiques), d'individus comme instance de concepts, et d'attributs. Les ontologies sont des modèles sémantiques de l'information très utilisés dans le courant actuel du Web 2.0, ou web sémantiques (Davies et al. 2002 ; Doan et al. 2002) : elles permettent ainsi de donner du sens à l'information partagée. Elles sont aussi très utilisées pour la description de méta-données en ingénierie des systèmes d'information (Guarino 1998).

Les travaux de Liu, Brewster et Shaw (Liu et al. 2013) présentent un état de l'art et une comparaison des ontologies proposées pour la gestion des crises, afin d'en identifier la structure commune. Ces travaux montrent qu'une grande partie de ces ontologies sont interopérables pour construire un lexique commun à tous types de crises durant toute la phase de gestion, grâce aux concepts de domaines tels que les ressources, les procédures, les organisations, les infrastructures, la météorologie, etc.

L'un des intérêts des ontologies est le développement de règles de raisonnement selon les concepts développés. Ces règles déterministes sont de la forme *SI -> ALORS*, et se fondent sur les relations modélisées au sein de la base de connaissances. Grâce à ces outils, de nombreux travaux en ingénierie des systèmes d'information ont pu améliorer l'utilisation de l'information par la sémantique. Les travaux de Roy ont porté sur la modélisation des types de navires et des évènements pour l'étude de leur comportement (Roy 2008 ; Roy 2009 ; Roy et Davenport 2009), pour l'amélioration des systèmes de surveillance. Dans cette continuité, les travaux de Vandecasteele se sont intéressés à la modélisation des mouvements et de la topologie des trajectoires, pour une classification automatisée des comportements des navires (Vandecasteele 2012 ; Vandecasteele et Napoli 2013 ; Vandecasteele et al. 2014). Ces travaux se basent notamment sur la mise au point de règles à partir des ontologies développées et de l'expertise de professionnels du domaine d'application.

L'utilisation de moteurs de règles peut aussi être choisie sans formaliser le contexte par des ontologies. Des exemples peuvent être trouvés dans la littérature sur des projets récents pour l'aide à la décision (Van Hille et al. 2012) ou pour la détection d'anomalies (Ray et al. 2013). Des règles spatiales, temporelles et/ou sémantiques permettent d'analyser les données d'un environnement choisi et comparer par rapport à des postulats émis. Les règles permettent d'orienter une décision, ou de mener une simulation afin d'anticiper la situation à venir, notamment grâce à des systèmes multi-agents (Murakami et al. 2002 ; Fournier 2005 ; Claramunt et al. 2007).

Dans le chapitre 2, nous reviendrons plus en détail sur la constitution de telles ontologies informatiques pour la gestion de connaissances. Nous présenterons notamment l'application d'un tel formalisme au domaine de l'information géographique. Toutefois, il est intéressant de noter que les règles utilisées par ce formalisme peuvent éventuellement être issues d'entretiens avec des experts, ou bien d'un processus de fouille de données qui aurait été mené en amont. La prochaine sous-partie présente la fouille de données (*data-mining*) qui peut être utilisée pour l'extraction de connaissances et de règles.

1.1.2.2.3. Fouille de données

La fouille de données, ou *data-mining* en anglais, peut être défini comme *une extraction non triviale de connaissances implicites et potentiellement utiles à partir des données* (Frawley et al. 1992). Cette technologie est utilisée afin d'extraire des connaissances au sein de grandes quantités de données, structurées dans des bases de données. Ces connaissances prennent la forme de motifs (*patterns*) reconnus de manière significative dans les données. Pour cela, les méthodes utilisées peuvent être automatisées ou semi-automatisées, selon le paramétrage souhaité par l'utilisateur. Le data-mining peut s'appliquer sur des données sémantiques classiques, ou bien sur des données spatio-temporelles (Roddick et Lees 2009).

Idiri (Idiri 2013 ; Idiri et Napoli 2012a) propose d'utiliser le data-mining pour la détection de routes maritimes, de zones « accidentogènes » (qui peuvent donner lieu à un accident) et de comportements anormaux de navires tels que des trajectoires aberrantes. Ce travail s'appuie sur des bases de données historiques sur les trajectoires et les accidents passés : l'analyse de plusieurs mois de données permet ainsi de dégager des scénarios d'intérêts, afin de mieux anticiper les risques depuis des données en temps réel. Un travail similaire est proposé par Maggi et al. (Maggi et al. 2013), grâce au *process mining* qui permet de découvrir, surveiller et améliorer des processus grâce à l'analyse de bases de données. La Figure 1-9 présente l'approche utilisée par ces deux études, pour l'analyse des comportements de navires.

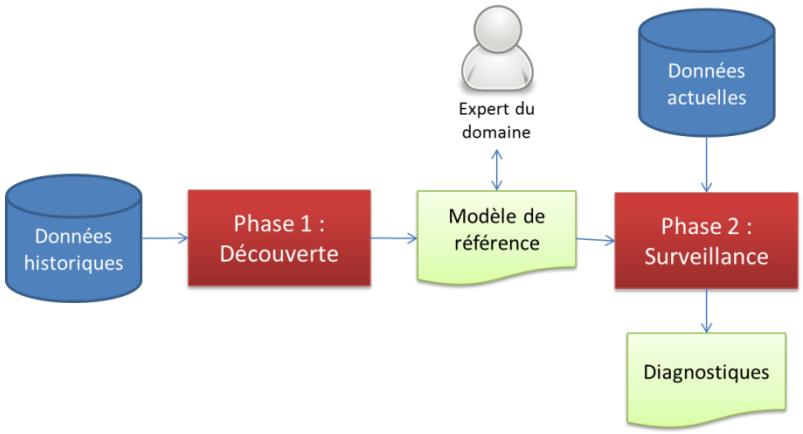


Figure 1-9. Analyse des comportements de navires par fouille de données (d'après Maggi et al. 2013)

Brisson, Collard et Pasquier (Brisson et al. 2006) proposent de coupler l'utilisation d'ontologie et de fouille de données pour le prétraitement de données et l'interprétation des résultats. Les ontologies servent ici à modéliser les connaissances des experts : cette modélisation de la connaissance est ensuite utilisée en aval de la fouille de données, afin d'interpréter les motifs découvert. Ces travaux montrent l'importance de la complémentarité entre ces nombreuses technologies : elles peuvent être utilisées séparément, mais la combinaison de leurs intérêts vient supporter la découverte et l'interprétation de résultats.

Cette interprétation automatisée des résultats de fouille de données se base sur la modélisation des connaissances expertes ; mais une autre manière de comprendre les résultats de fouille de données peut être obtenue par la visualisation, interprétée par des experts.

1.1.2.2.4. Visualisation

Dans une partie précédente de ce manuscrit, nous avons introduit la notion de *Situation Awareness*, qui fait appel, entre autres, à l'analyse, l'interprétation d'information ainsi que leur projection dans le futur proche (Endsley 1995 ; Glandrup 2013). L'être humain est au cœur de ce processus d'analyse et de décision, où il interagit avec la machine et sa puissance de calcul. Pour cela, l'opérateur doit donc nécessairement interagir avec la machine, notamment par la visualisation d'information. La visualisation lui permet ainsi de valider une hypothèse (analyse) ou bien de mener l'exploration elle-même au sein de données (Keim et al. 2010 ; Thomas et Cook 2005). Au-delà de l'affichage, la visualisation a aujourd'hui évolué vers l'analyse visuelle, qui permet de coupler la puissance des technologies identifiées précédemment avec des interfaces hautement interactives (Keim et al. 2010).

Dans le domaine de la gestion des risques et des crises, Akella (Akella 2009) propose une étude des icônes utilisées sur des cartes de risques avec une équipe de pompiers, afin d'évaluer la compréhension et l'efficacité de ces symboles. Son étude montre que l'utilisation de symboles qui semble clairs, peut parfois avoir une signification inattendue pour les utilisateurs finaux. Avant de souligner l'importance de la visualisation pour l'analyse des données pour l'analyse d'une situation, cette étude souligne le rôle premier de la visualisation, qui est la **communication** d'informations. Cette étape est en effet le premier niveau de la visualisation d'information selon MacEachren (MacEachren 1995), théorie sur laquelle nous reviendrons précisément dans le chapitre 2.

Tomaszewski et al. (Tomaszewski et al. 2007) montrent le rôle important de la visualisation d'information géographique pour l'analyse, la gestion des risques et la prise de décision, dans de nombreux cas tels que les catastrophes naturelles, la santé, etc. La visualisation tient une place fondamentale dans les phases de communication d'information, d'analyse post-évènement, d'étude de la vulnérabilité ou pour l'exercice à la gestion de crises. Nous retrouvons ainsi des grands axes communs avec les buts de la SA : extraction d'information, modélisation et anticipation (Tomaszewski et al. 2007 ; Andrienko et Andrienko 2007 ; Seppänen et Virrantaus 2010 ; Cai 2005).

Dans le domaine du maritime, Davenport et Risley (Davenport et Risley 2006) proposent un état de l'art des différentes méthodes de visualisation de l'information, qui permettent de supporter l'analyse et la prise de décision. Les auteurs comparent les possibilités offertes par les très nombreuses méthodes et logiciels de visualisation d'information, appliqués à l'étude des trajectoires de navires. Les travaux de Riveiro (Riveiro et al. 2008 ; Riveiro et Falkman 2011) mettent aussi en avant l'importance de la visualisation comme moyen d'aide à la prise de décision, dans une domaine aussi sensible que le trafic maritime.

La puissance du data-mining et l'interactivité de la visualisation peuvent aussi être combinés afin d'aider l'analyse d'informations et la prise de décision, au sein de nouvelles technologies intelligentes. La dernière sous-partie traite d'une solution technologique récemment appliquée au domaine spatial : OLAP et SOLAP.

1.1.2.2.5. OLAP et SOLAP

L'acronyme OLAP signifie *Online Analytical Process* (Codd et al. 1993 ; Chaudhuri et Dayal 1997 ; Berson et Smith 1997), soit traitement analytique en ligne. Cette technologie est très utilisée dans le domaine de l'information décisionnelle (Abdullah 2009), afin d'étudier de grandes quantités de

données dont l'une des dimensions est le temps. La technologie OLAP permet d'organiser l'information en une structure spécifique dans des bases de données (structure en étoile, structure en flocon), selon plusieurs dimensions temporelles ou attributaires.

Les travaux de Bédard ont introduit le concept d'OLAP spatial, ou SOLAP, pour l'étude de données à composantes spatiales à plusieurs niveaux (Bédard et al. 2007 ; Rivest et al. 2001). Les auteurs mettent en avant la différence d'une approche SOLAP par rapport à l'approche classique SIG pour la prise de décision dans une étude de phénomène spatial, grâce à la rapidité et la facilité d'utilisation du SOLAP, ainsi que le niveau de détail permis (Figure 1-10).

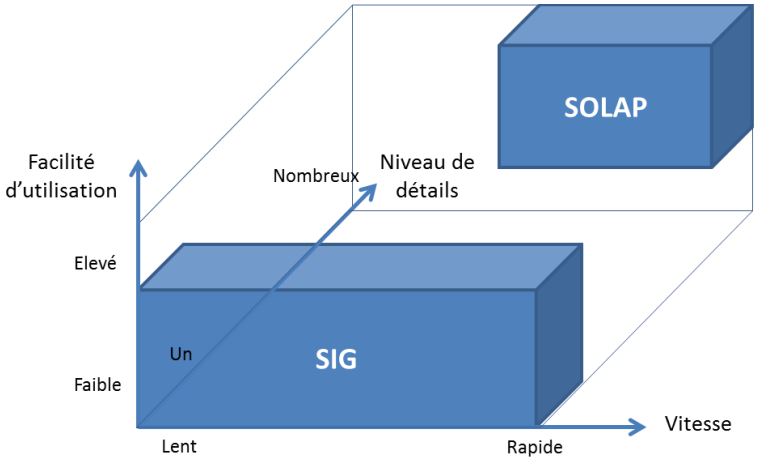


Figure 1-10. Comparaison des SIG et du SOLAP pour l'aide à la décision (d'après Rivest et al. 2001)

Les données OLAP sont modélisées sous forme de cube de données, qui permet ensuite des opérations de fouille telles que le *drill-up / drill-down*, *drill-through*, *rotate*, etc. (voir Figure 1-11). Les différents niveaux des données étudiées sont ainsi approfondis ou généralisés, inter-changés, afin d'évaluer les différents liens entre les données par des visualisations de l'information. Ces visualisations prennent généralement la forme de tableaux, graphes ou cartes choroplèthes qui peuvent être utilisés de manière interactive (Beaulieu 2009).

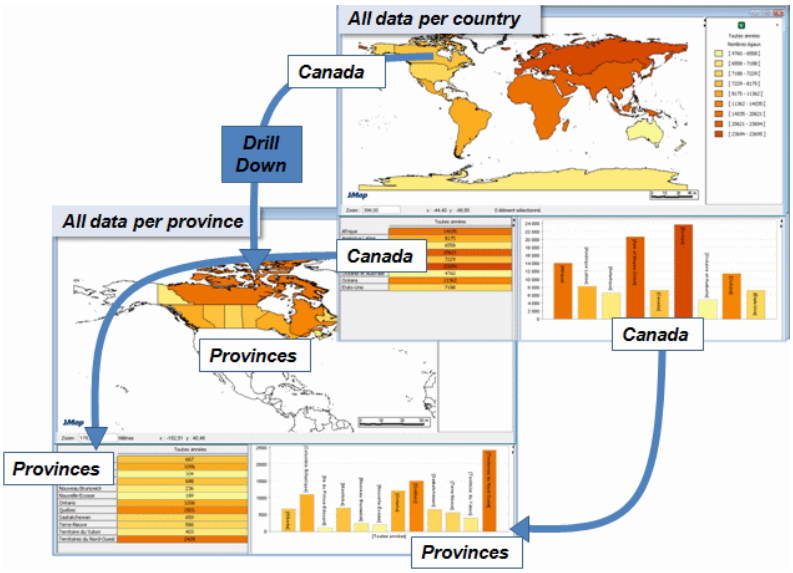


Figure 1-11. Opération de drill-down en SOLAP (*source : spatialolap.scg.ulaval.ca*)

L'avantage de la technologie OLAP est de générer ces statistiques en temps réel (en ligne), ne demandant donc pas de long temps de calcul. Néanmoins, l'OLAP nécessite une lourde phase de modélisation de l'information, de structuration des tables en vue de la création d'un cube de données, ainsi qu'une phase d'intégration des données brutes vers ce cube. L'évolution des données demande donc un important travail de mise à jour de la structure OLAP.

Dans cette première partie, nous avons donc pu voir les nombreux processus possibles en gestion de risques, et les diverses technologies utilisées pour supporter l'analyse de situation et la prise de décision. Ces technologies, qu'elles soient orientées traitement de l'information automatisé ou visualisation de l'information par l'être humain, permettent de mettre au point de nouveaux outils pour améliorer la gestion des risques du domaine étudié. Toutefois, dans un tel monde opérationnel, les utilisateurs tels que contrôleurs ou décideurs sont généralement habitués à une certaine classe d'outils, de systèmes, qu'ils utilisent dans l'analyse des données (Enguehard et al. 2013b ; Chuttur 2009 ; Davenport et Risley 2006 ; Riveiro 2011 ; van Laere et Nilsson 2009). Les outils issus du monde de la recherche se retrouvent souvent dans une situation complexe, où l'utilisateur final n'est pas directement prêt à faire usage de cette nouvelle technologie : ceci peut venir d'un manque de formation, et ainsi d'une mauvaise connaissances des données et solutions nécessaires à manipuler les outils d'exploration (MacEachren 1995), ou bien une non concordance de la proposition et des besoins réels, soit le fameux décalage entre la recherche théorique et l'application (McCollum 2007). Dans la prochaine partie de ce chapitre, nous présentons les éléments qui peuvent freiner l'acceptation de nouvelles technologies au sein d'un monde opérationnel.

1.2. L'utilisateur et l'usage des nouvelles technologies

Dans de nombreux domaines, nous pouvons observer que les systèmes de contrôle sont encore limités dans leur utilisation, car des outils d'analyse des informations ne sont que rarement proposés (Glandrup 2013). Cela est souvent dû à l'inéquation entre les outils utilisés et leur réelle finalité d'utilisation, la prise de décision. Pour l'utilisation de données géographiques et la prise de décision, de nombreux outils du type SIG sont utilisés. Pourtant, ces outils sont basés sur l'affichage et la cartographie de données, pour une utilisation effectuée par un unique individu (MacEachren et Brewer 2004). Selon MacEachren et Brewer, les SIG ne sont donc pas adaptés au processus de prise de décision, qui se fonde sur la collaboration de nombreux acteurs et une analyse approfondie des données. Pour cette problématique socio-technique, le concept de géocollaboration est alors introduit. Il représente l'interaction entre les nombreux acteurs dans la gestion de risques ou de crises, basé sur l'utilisation de systèmes communs pour l'analyse de données spatiales (Jankowski et al. 1997 ; MacEachren 2001).

Pourtant, dans le domaine de la surveillance et du contrôle, l'analyse visuelle et les interfaces utilisées jouent un rôle crucial dans la compréhension de l'information et des décisions qui sont prises en aval. C'est cette étape de visualisation de l'information qui fait la transition entre les données concrètes et l'interprétation que l'on peut en faire en en faisant de l'information, puis des connaissances (Skourup et Aune 2001).

> *Operators do not need more data, they require more information, which means that data is put into a context such as in an advanced information presentation. (Skourup et Aune 2001)*

L'introduction des concepts de (géo-) collaboration et de « présentation avancée de l'information » mettent donc en évidence les limites d'usage des outils « classiques » de visualisation et de prises de décision. Ces limites pour le contrôle et la prise de décisions, qui portent notamment sur l'utilisation de la visualisation d'information géographiques, ont de nombreuses causes. Nous pouvons citer, entre autres :

- Les données utilisées : contenu, qualité, structure, volume, présentation, etc. Pour cela, une étude complète a été menée par Harding, afin d'identifier les facteurs qualitatif pour l'usabilité de l'information géographique (Harding 2013).
- Des limites propres à l'être humain, telles que les limites cognitives, visuelles, suite à une mauvaise représentation de l'information ou de mauvais choix technologiques (Harrower 2007)
- L'acceptation de nouvelles technologies, qui peut être freinée par l'organisation actuelle d'une communauté : outils, acteurs, habitudes et bonnes pratiques (Davis et al. 1989 ; Davis 1993 ; Chuttur 2009)
- La structure organisationnelle de la décision et l'utilisation coopérative des outils pour la décision (Noucher 2013 ; Chemitte 2008)

Dans la suite de cette partie, nous revenons particulièrement sur les limites de la cognition humaine pour le contrôle, ainsi que l'acceptation des technologies. Il s'agit en effet de deux axes majeurs afin d'étudier les solutions existantes pour le contrôle et proposer leur utilisation adaptée au contexte.

1.2.1. Les limites de la cognition humaine face à la visualisation

L'utilisation d'interface hommes-machines pour l'étude et le contrôle d'information nécessite de porter une grande attention sur la manière dont fonctionne la pensée humaine, notamment la pensée visuelle. Par le fait de surveiller un écran, qui affiche des informations d'intérêt, la pensée visuelle permet de raisonner par les images que nous percevons, et ce de manière rapide (MacEachren 1995 ; Ware 2008). Par la visualisation, l'être humain peut extraire plusieurs motifs et relations entre les éléments visualisés, plus facilement qu'il ne peut le faire depuis de grandes quantités de texte par exemple.

Toutefois, le mode de fonctionnement de notre système visuel ne nous permet pas de percevoir l'intégralité de l'information affichée et des relations qui peuvent exister au sein de ces données. Tout d'abord, la limite principale provient de la surcharge de l'information affichée, par laquelle notre perception visuelle est rapidement saturée (Miller 1956). Miller met en avant la distinction d'un certain nombre de variables, ou classes, à partir duquel nous n'arrivons plus à distinguer les différences entre ces nombreux éléments : dans son article original, le chiffre donné est de **7, plus ou moins 2**. Nous retrouvons ce chiffre dans l'outil cartographique ColorBrewer (Harrower et Brewer 2003) par exemple, où un maximum de 9 valeurs de couleurs sont proposées pour distinguer des classes séquentielles (voir Figure 1-12). Cette valeur de 7, qui a récemment été identifiée à la baisse pour une valeur de 3 (Farrington 2011), représente donc le maximum d'informations que peut contenir notre mémoire de travail, ou mémoire opérationnelle, qui conserve les informations le temps de la réalisation d'une tâche (Richard 1982).

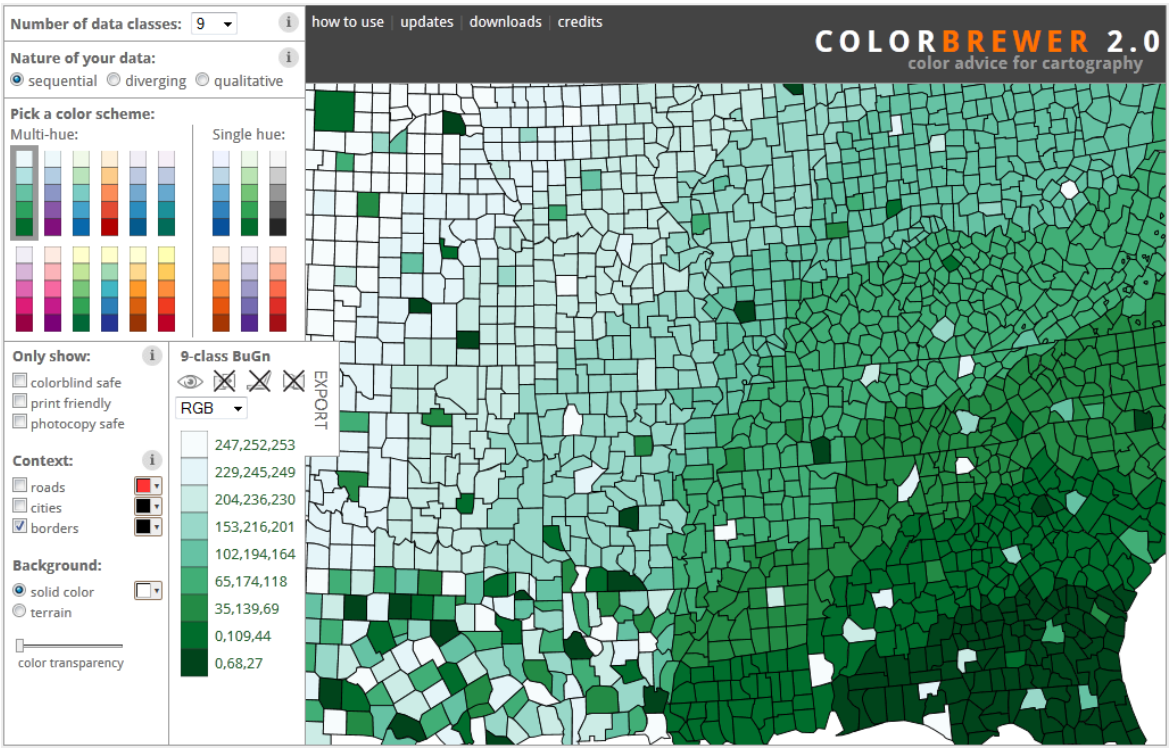


Figure 1-12. ColorBrewer 2.0, proposition d'échelles de couleurs selon la classification des données

La surcharge d'information ne se retrouve pas seulement lors de l'affichage de données statiques : cette limite s'observe aussi lors de l'affichage de données dynamiques, ou animées. Par exemple, dans le domaine de la cartographie, l'utilisation de cartes animées pose un grand problème de compréhension et d'interprétation de la donnée (Wilkening et Fabrikant 2011 ; Harrower 2007). En effet, l'affichage de cartes choroplèthes successives afin d'illustrer l'évolution d'un phénomène, est un bon exemple de difficulté à interpréter les informations. Si le nombre d'entités et l'étendue spatiale sont trop élevés, l'utilisateur peut avoir du mal à identifier les changements et à se souvenir de la valeur passée d'un objet. Les paramètres sur lesquels jouer afin d'améliorer une telle animation serait les contrôles d'une telle interface, le temps d'affichage, le nombre de classes, etc.

La perception de l'information dépend aussi fortement du but d'utilisation, qui aurait été identifié en amont. En effet, notre perception se focalise sur certains objets, suivant la tâche qu'il y a à accomplir : on parle alors d'attention sélective. Cette attention biaise fortement notre perception de l'information, car des objets particuliers sont alors recherchés. C'est le cas dans l'expérience du « Gorille Invisible » (Simons et Chabris 1999), qui montre par la vidéo une erreur importante (élément *absurde* dans la scène, non repéré) due à notre attention sélective.

Ce processus d'attention sélective met en évidence une certaine manière de fonctionner de la perception visuelle, à savoir son approche *top-down* qui est centrée sur les concepts à rechercher. Une approche inverse est dite *bottom-up*, où depuis la visualisation d'un ensemble, nous cherchons à extraite une information particulière : il s'agit alors de vision pré-attentive (Ware 2008).

Nous aurons l'occasion de revenir sur les différentes manières de tirer profit de ces erreurs de perception, lorsque nous décrirons plus en détail la visualisation dans le chapitre 2. Ce paragraphe a permis d'identifier les limites majeures que notre perception peut poser, lorsque nous devons étudier des informations visuelles. En situation de contrôle et de prise de décision, de telles limites

doivent évidemment être connues et maitrisées, afin de ne pas manquer des éléments importants pour la décision, ou bien afin de ne pas comprendre la mauvaise information.

1.2.2. L'acceptation de nouvelles technologies

Dans les domaines de la surveillance et du contrôle, de très nombreux outils sont encore proposés chaque année par de nombreux laboratoires de recherche ou industriels. Parmi ces nouvelles technologies pour l'analyse et la surveillance, nous pouvons distinguer trois grandes catégories :

- Les technologies automatisées
- Les technologies semi-automatisées
- Les technologies basées sur la visualisation

Les technologies automatisées privilégient le calcul machine dans de très grandes quantités de données, afin d'en extraire automatiquement de nouvelles connaissances et reconnaitre des motifs d'intérêt. Cette automatisation est proche des problématiques soulevées par le courant du *Big data*, *i.e.* l'enregistrement, le stockage, le partage et l'analyse de très grandes quantités de données. Nous pouvons citer l'utilisation de la fouille de données (*data mining*) pour la recherche de motifs dans de grandes bases de données structurées (Fayyad et al. 1996 ; Cabena et al. 1997 ; Han et Kamber 2006). De nombreux travaux ont été réalisés afin d'étudier l'apport de la fouille de données à l'étude de données de mouvement, et passer ainsi d'un ensemble de données à une collection d'information et de scénarios récurrents, ou bien anormaux (Idiri 2013 ; Idiri et Napoli 2012a ; Etienne 2011).

L'utilisation de technologies semi-automatisées permet d'inclure l'être humain dans le contrôle des étapes de l'analyse de données, afin de régler plusieurs paramètres et raffiner les résultats issus d'un processus automatisé. Des algorithmes de fouille de données (*data mining*) permettent par exemple de maitriser et améliorer les résultats de recherches, selon les choix de l'utilisateur. Enguehard et al. (Enguehard et al. 2013a) comparent ainsi l'utilisation de plusieurs outils, pour une même tâche d'analyse de comportement de navires de pêche, en fondant cette approche sur des solutions automatisées, semi-automatisées ou visuelles. La solution semi-automatisée se base sur une approche automatisée, mais dont l'utilisateur peut consulter, modifier les paramètres, jusqu'à atteindre des résultats satisfaisants.

Mais face à de trop nombreuses propositions d'outils automatisés pour l'analyse de données et la prise de décisions, se pose alors la question de la légitimité des résultats et des décisions issus de tels systèmes. En effet, une automatisation trop importante du processus d'analyse met alors l'humain et sa propre expérience au second plan, dans la prise de décisions. Pour en étudier l'impact en termes d'usage et de confiance de tels systèmes, nous présentons plusieurs travaux et réflexions qui ont été menés de manière à évaluer la pertinence de tels outils.

1.2.2.1. Automatisation / Interaction : limites ou complémentarité pour l'analyse d'information ?

Comme nous l'avons vu depuis l'introduction de ce manuscrit, différents axes sont suivis par les chercheurs afin de proposer de nouvelles solutions d'analyse de l'information, pour la prise de décision. Nous avons identifié des travaux menés sur une analyse automatisée des données, semi-automatisées, ou bien une analyse entièrement menée par des opérateurs humains grâce à la visualisation. Face à ces nombreux travaux, la question se pose donc de savoir quel serait l'équilibre

idéal entre une analyse entièrement menée par une machine, par automatisation, et une recherche entièrement menée par un ou des opérateurs humains. Nous présentons ici quelques travaux de chercheurs qui se sont intéressés à cette question.

Un premier positionnement à considérer sur ce problème est la vision donnée par Parasumaran, Sheridan et Wickens (2000) pour étudier le « niveau d'automatisation » de systèmes pour la prise de décision. Tout d'abord, les auteurs identifient quatre étapes majeures dans le processus de traitement de l'information par un être humain, en vue d'une prise de décision. Ces étapes sont alors, dans l'ordre :

- Acquisition d'information : capture et enregistrement de l'information à analyser, grâce à une combinaison spécifique de capteurs. Des opérations de filtrage de l'information peuvent aussi intervenir dès cette étape.
- Analyse de l'information : prédiction des données à venir, intégration des différentes sources d'information, détection de motifs, etc. L'automatisation permet donc de pré-traiter l'information et la présenter de manière à faciliter sa compréhension par l'humain (amélioration de la perception et cognition).
- Sélection de la décision : selon les résultats de l'analyse menée et une liste de décisions possibles, le système doit pouvoir sélectionner la décision la plus adaptée à la situation.
- Implémentation de l'action : la décision sélectionnée doit être appliquée, par l'humain ou la machine, en attendant une certaine réponse des données observées.

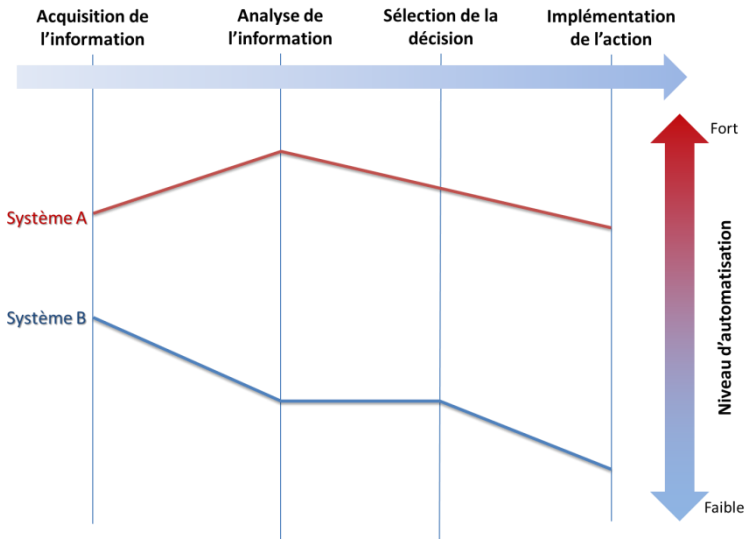


Figure 1-13. Différents niveaux d'automatisation sur les fonctions lors de la prise de décision (d'après Parasuraman et al. 2000)

La Figure 1-13 illustre le modèle proposé par Parasumaran, Sheridan et Wickens, afin de décrire le niveau d'automation selon les étapes successives de la gestion de l'information et de la prise de décision. Les auteurs proposent d'évaluer la portée de chacun de ces quatre axes sur la performance humaine, avant de décider du niveau d'automatisation à autoriser pour un système donné. Plusieurs conséquences négatives sur l'opérateur ont ainsi été identifiées, suite à une utilisation abusive de l'automatisation : parmi celles-ci, une compréhension de la situation dégradée, une autosatisfaction ressentie par l'opérateur, ou encore une perte de compétences (Parasuraman et al. 2000, 291).

L'automatisation de tâches pour l'analyse d'informations et pour la prise de décision est donc un sujet sensible, qu'il ne faut pas séparer du processus mené par l'humain lui-même : pour cela, la comparaison des gains et pertes entre un jugement humain et une analyse automatisée est un axe de recherche majeur depuis les années 90 (Bass et Pritchett 2008 ; Parasuraman et Riley 1997). Bass propose la méthode HAJL (*Human-Automated Judge Learning*) pour entrainer l'homme à interagir avec un processus automatisé lors de la prise de décision (Bass et Pritchett 2008 ; Bass et al. 2013). Cette méthode est fondée sur trois phases successives : l'entrainement, l'apprentissage et la prédiction. Elle permet de mesurer des indicateurs d'interaction entre l'homme et l'automatisation, notamment l'acceptation des méthodes automatisées, la similarité entre les décisions humaines et automatiques, ou encore le degré de réussite estimé par l'humain seul. La méthode HAJL met en avant la complémentarité entre l'automatisation et l'analyse menée par l'humain dans la prise de décision.

De manière proche, les travaux menés par Enguehard et al. permettent d'identifier les limites d'outils trop automatisés pour l'analyse et le contrôle d'une situation, notamment pour l'étude des mouvements de navires de pêche (Enguehard et al. 2013b ; Enguehard et al. 2013a ; Enguehard et al. 2012). Selon les auteurs, l'automatisation et l'interactivité au sein d'outils informatiques ne sont pas deux aspects distincts, mais complémentaires. Un espace à deux axes est donc proposé afin de positionner les différents environnements évalués, selon le niveau automatisation et le niveau d'interactivité (voir Figure 1-14). Nous pouvons assimiler cette approche à une synthèse du modèle proposé par Parasumaran, Sheridan et Wickens : un système peut être à la fois très automatisé et très interactif, mais **à différents niveaux dans le processus de traitement des données** (les quatre axes de la Figure 1-13).

Enguehard et al. appliquent leur étude à plusieurs systèmes pour l'étude des comportements de navires de pêche, testés et évalués par des professionnels de la surveillance maritime (Enguehard et al. 2013a) : certains sont entièrement automatisés, tandis que d'autres proposent une interaction avec les données et les modèles de comportements (par exemple, changer les classes identifiées automatiquement). Sur la Figure 1-14, nous présentons de manière générale les outils qui étudiés par Enguehard, ainsi que les positions approximatives d'autres méthodes qui permettent de visualiser les mouvements de navires, à titre d'exemple. Le système HSF (*Hybrid Spatio-temporal Filtering*) est présenté par l'auteur comme nouvelle manière de combiner la puissance du calcul machine et du raisonnement humain pour la classification des comportements de navire (Enguehard et al. 2013b). Sur cette figure, nous identifions trois groupes de méthodes pour analyser de tels types de mouvements : les systèmes de surveillance utilisés actuellement dans les centres de contrôle (affichage de données, très peu de contrôle laissé à l'utilisateur), les projets R&D menés pour la surveillance maritime (intelligence artificielle, règles) et le domaine de l'analyse géovisuelle qui, par définition, propose un haut niveau d'interaction couplé à la puissance de calcul machine (différents niveaux d'automatisation).

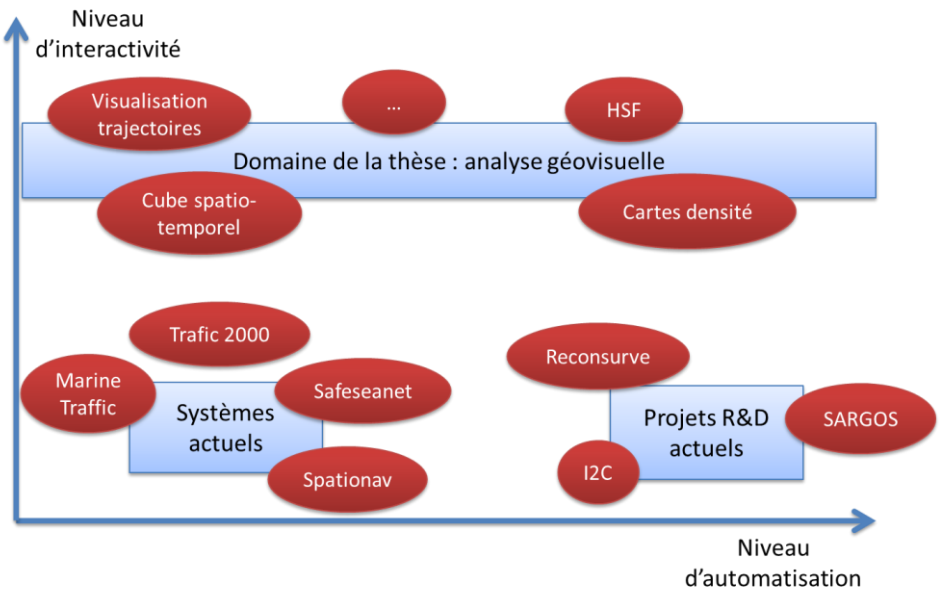


Figure 1-14. Interactivité / Automatisation : deux aspects complémentaires pour l'étude de données spatiales pour l'analyse des risques maritimes. Les exemples sont placés de manière approximative.

La question de l'automatisation est d'autant plus importante en situation de crise, ou en situation où le temps fait pression : les données doivent être mises dans le contexte et analysées de la manière la plus rapide possible, tout en conservant une précision optimale. Cela fait appel au concept de *Situation Awareness*, décrit dans la partie 1.1.2. Luokkala et Virrantaus (Luokkala et Virrantaus 2014) montrent qu'en situation de crise, la SA active deux processus mentaux parallèles : le premier est inconscient et basé sur les réflexes, appelé « système 1 » ; le second est réfléchi et conscient, appelé « système 2 ». Le premier système mental autorise la reconnaissance automatique de motifs dans les données, tandis que le système 2 nécessite d'estimer les différents scénarios et solutions pour trouver une solution raisonnée. Face à ces deux systèmes complémentaires, l'humain est influencé par des facteurs **structurels** (expérience, formation, intérêts personnels, etc.) et **situationnels** (mission, conditions de travail, etc.) (Nofi 2000 ; Luokkala et Virrantaus 2014). Ces deux systèmes mentaux et les nombreux facteurs qui les influencent demandent donc différentes méthodes pour améliorer la SA d'un environnement donné. Les auteurs donnent comme exemples d'outils le data-mining géographique ou bien l'utilisation de cartes pour supporter le système 1 (reconnaissance rapide des motifs), tandis que le système 2 pourrait être supporté par un processus plus « narratif » pour analyser l'information (*story-telling*) (Lundblad 2013 ; Ho 2013 ; Ho et Jern 2013 ; Robinson 2008).

Suite à ces travaux, nous pouvons comprendre que les niveaux d'automatisation et d'interactivité, ou de contrôle par l'humain lui-même, ne sont pas si dissociables lors de l'analyse d'informations. Ceci est encore plus valable lors de la gestion d'évènements tels que les crises, les dangers imminents, où le temps fait pression sur la prise de décision. Dans le dernier paragraphe de cette partie, nous avons introduits certains facteurs qui ont une influence importante sur la prise de décision et l'utilisation de certaines technologies. Dans la partie suivante, nous revenons plus en détails sur ces différents freins socio-culturels pour l'utilisation de nouvelles technologies.

1.2.2.2. Freins socio-culturels pour l'acceptation des technologies

1.2.2.2.1. TAM : le modèle de l'acceptation de la technologie

Lorsque l'on parle d'innovation et de nouvelles technologies, il est très courant d'observer une réticence des utilisateurs visés lors de l'intégration de ces nouveaux outils dans leur procédure habituelle. De même, il existe encore une marge importante dans l'utilisation effective de solutions développées par la recherche, et les potentiels utilisateurs de ces nouvelles solutions. Dans cette partie, nous présentons les travaux qui ont pu être menés récemment sur l'acceptation de nouvelles technologies, afin de comprendre comment proposer une innovation qui soit utile et utilisée de manière effective dans le système amélioré.

Les nombreuses études de Davis ont amené plusieurs modèles successifs pour décrire les facteurs influençant l'acceptation de la technologie, le *Technology Acceptance Model* ou TAM (Davis et al. 1989 ; Davis 1993 ; Benbasat et Barki 2007 ; Chuttur 2009 ; Bagozzi 2007). Deux facteurs décisifs sont identifiés par Davis et son équipe (Davis et al. 1989), à savoir l'**utilité perçue PU (*Perceived Usefulness*)** et la **facilité d'utilisation perçue PEOU (*Perceived Ease-Of-Use*)**. Par la suite, nous utilisons les acronymes anglo-saxons pour parler de ces facteurs majeurs.

Le modèle original du TAM remonte à 1989 (Davis et al. 1989, 19), mais des premières corrections ont été amenées pour aboutir au modèle le plus reconnu quelques années plus tard (Davis 1993) (voir Figure 1-15). Les deux facteurs PU et PEOU sont définis comme le ressenti personnel de l'utilisateur face à de nouveaux outils pour les tâches qu'il doit effectuer avec le système. En effet, il est important de différencier la perception d'un utilisateur par rapport à l'attitude et le comportement réel (Fishbein et Ajzen 1975 ; Davis 1993). Pour cela, Davis définit ainsi ces deux facteurs cognitifs :

- L'utilité perçue (PU) est l'évaluation selon laquelle un système permettra d'améliorer les tâches habituelles de l'utilisateur.
- La facilité d'utilisation perçue (PEOU) est la perception, du point de vue d'un utilisateur, qu'un système ne demandera pas ou peu d'effort pour être utilisé.

Comme illustré sur la Figure 1-14, la PU dépend de la PEOU : si un outil n'est pas vu comme simple à utiliser, l'évaluation de son utilité en sera alors amoindrie. La facilité d'utilisation est donc un facteur clé dans l'utilisation effective de nouveaux outils. De ces deux facteurs dépend alors la première impression de l'utilisateur potentiel, qui présente alors son intention d'utiliser ou non le nouvel outil. Au final, cette nouvelle technologie proposée sera, ou non, utilisée de manière effective par les différents opérateurs, selon ces nombreux facteurs.

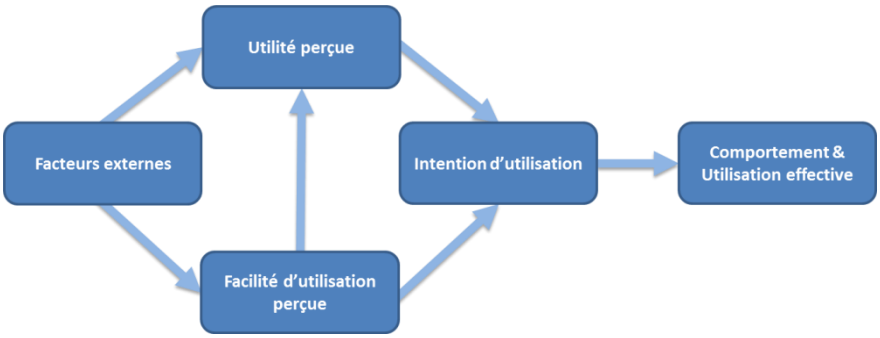


Figure 1-15. TAM : Modèle de l'acceptation de la technologie corrigé (d'après Davis 1993)

Le modèle TAM a ensuite été repris et étendu de nombreuses fois, de façon à approfondir la compréhension des facteurs externes qui influent sur la PU et la PEOU (Venkatesh et Davis 2000 ; Venkatesh et al. 2003 ; Venkatesh et Bala 2008). TAM 2 apporte alors plus de détails sur les facteurs externes qui jouent sur l'utilité perçue d'un nouvel outil (voir Figure 1-16). Le modèle le plus récent, proposé par Venkatesh et Bala (Venkatesh et Bala 2008), porte notamment sur les facteurs fondamentaux qui influent sur l'acceptation du commerce en ligne (*e-commerce*). Le modèle TAM 3 a alors l'intérêt d'exploiter les facteurs qui jouent sur la facilité d'utilisation perçue et qui peuvent être partagés en deux catégories : les ancres qui influent négativement, et les ajustements qui viennent soutenir la PEOU (voir Figure 1-17).

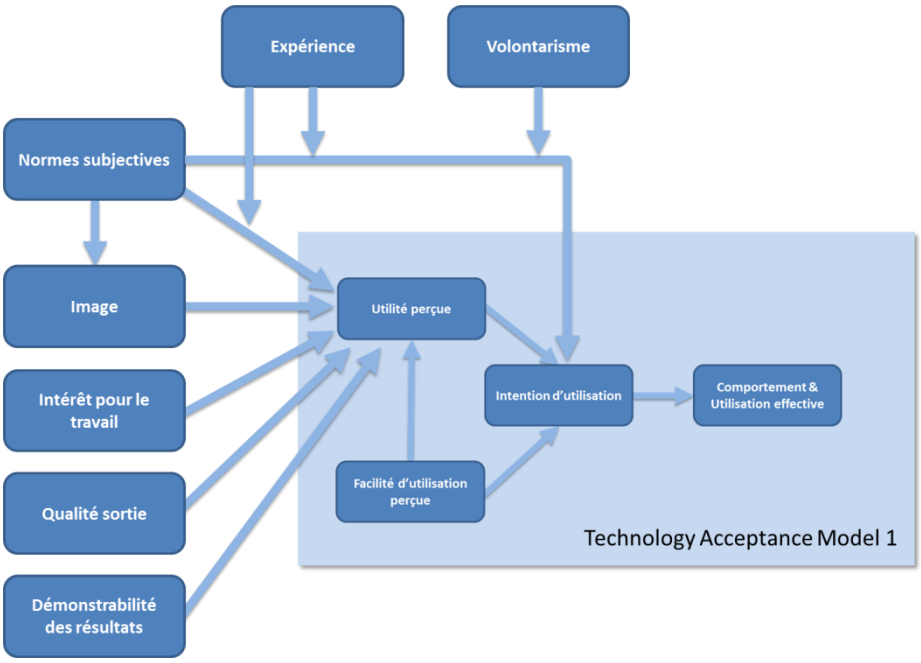


Figure 1-16. Modèle TAM 2, décrivant les facteurs d'influence sur l'utilité perçue

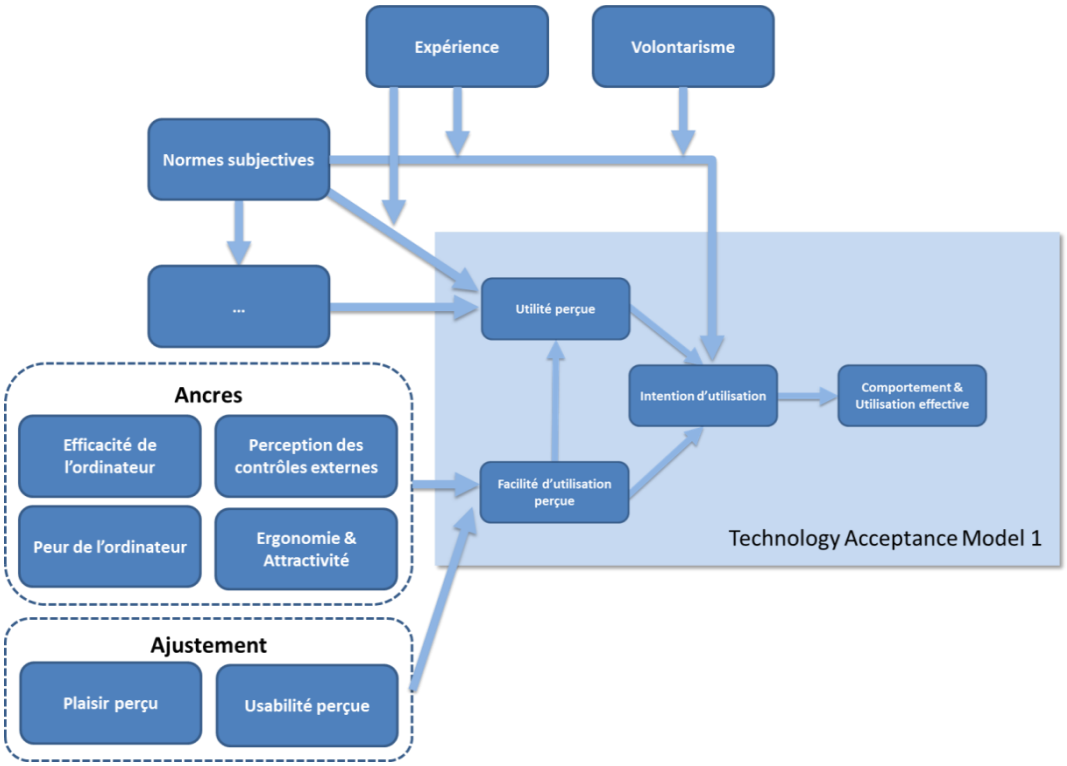


Figure 1-17. Modèle TAM 3, réutilisant les facteurs de TAM 2 (non placés ici) et les facteurs d'influence sur la facilité d'utilisation perçue

Mais de récentes critiques ont été faites sur l'utilisation abusive du modèle TAM, le manque d'innovation des chercheurs de ce domaine, et la confiance trop grande qui a pu être portée sans faire place à l'innovation (Benbasat et Barki 2007). Bagozzi (Bagozzi 2007) a suggéré que la mémoire humaine puisse ne pas être influencée par les facteurs mis en avant par le TAM, créant ainsi un biais important. Ainsi, l'auteur met en avant la fragilité du lien entre l'intention et le comportement, où la période de temps entre les deux peut induire de nombreux autres facteurs d'incertitude. Avant tout, les critiques du TAM reprochent le déterminisme de ce modèle (Bagozzi 2007 ; Chuttur 2009). Le TAM est donc un modèle qui a été innovant en son temps, qui a pu servir de base importante pour le travail de plusieurs groupes de recherche. Mais de nombreuses recherches se sont trop appuyées sur le TAM comme acquis, ce qui a pu faire oublier de mener une recherche fondamentale sur les facteurs et leurs liens quant à l'acceptabilité d'outils nouveaux.

A l'échelle de notre recherche, le TAM reste néanmoins une base d'intérêt dans l'optique de proposer de nouveaux outils de contrôle et d'analyse de comportements : pour cela, nous retenons donc les facteurs PU et PEOU, ainsi que d'autres facteurs externes qui seront utiles pour notre étude menée.

1.2.2.2.2. Limites des systèmes d'information pour la prise de décision

La prise de décision est issue du travail de plusieurs personnes, amenées à collaborer sur un travail d'analyse sous une pression potentielle (urgence, gestion de crise immédiate, etc.). Pour cela, en plus des limites (cognitives, sociales, culturelles, etc.) d'un utilisateur seul face à un système, il faut prendre en compte l'inertie d'un groupe entier pour la prise de décision. Des compromis doivent être éclaircis sur les questions des méthodes, des acteurs et des outils (Noucher 2013). Toutes ces ressources sont mobilisées pour collaborer dans l'analyse de la situation et la prise de décision. Au

sein de cette collaboration, il est donc nécessaire de prendre en compte de nombreux facteurs déterminants pour l'adoption de méthodes et de données, qui peuvent être d'ordre sociologiques, culturels ou cognitifs (Noucher 2013 ; Caron et Roche 2004).

Pour le cas de l'utilisation d'information géographique dans un cadre collaboratif, Noucher (Noucher 2013 ; Noucher 2009) propose de voir ces données comme un « *objet-frontière cognitif et collaboratif* » qui permet un dialogue entre les différents profils d'acteurs. Il met alors en évidence la nécessité de voir l'information géographique comme un objet cognitif et sémantique, plutôt qu'une donnée simple. Dans un cadre de collaboration, il est donc important de voir les données utilisées par leur sens réel (sémantique) et non pas l'information qu'elle porte. Ainsi, il initie une nouvelle façon de modéliser et d'utiliser l'information au sein de systèmes d'aide à la décision, qui diffèrent fortement de systèmes d'information classiques. En effet, les systèmes d'information tels que les SIG ont été pensés pour une utilisation par une personne seule (MacEachren et Brewer 2004 ; Cai 2005), contrairement aux systèmes d'aide à la décision qui doivent être adaptés à une utilisation collaborative (Wu et al. 2009 ; MacEachren et Brewer 2004 ; Seppänen et Virrantaus 2010).

En plus des limites posées par les méthodes, les acteurs ou les outils, l'acceptation de nouvelles méthodes peut aussi dépendre de la qualité des informations elles-mêmes, utilisées pour l'analyse et la prise de décision. Pour cela, Harding (Harding 2013) identifie et évalue neuf catégories de facteurs qualitatifs qui influent sur l'usabilité d'information géographique, et donc l'acceptation (d'après le modèle TAM) de nouvelles technologies. Cette étude peut facilement être appliquée à des informations autres que géographiques (Harding 2013), car ces facteurs ne sont pas spécifiques à des données spatiales. Ceux-ci sont résumés sur la Table 1-2, basée sur l'article de Harding. Ce tableau donne la liste des neuf facteurs identifiés, une courte description de ces concepts, ainsi que le possible impact négatif qu'ils peuvent avoir sur l'analyse et l'utilisation de l'information dans le système. La dernière ligne introduit le concept de « sélectabilité » (*selectability*) qui correspond à la possibilité de filtrer les données de manière à réduire la charge d'informations, selon l'utilisation souhaitée du système. L'auteur met alors en avant les avancées du web sémantique pour améliorer l'intégration des différentes sources de données et pour renseigner sur ces nombreux facteurs.

Parmi ces facteurs, les entretiens menés par Harding montrent un souci prononcé pour le contenu des données, la qualité des données et la structure des données. Le contenu des données est notamment la première préoccupation des groupes d'utilisateurs pour l'analyse, la modélisation, la gestion de risques et la gestion de trafic.

Table 1-2. Facteurs qualitatifs pour l'usabilité de l'information géographique (d'après Harding 2013)

Facteur	Description	Impact négatifs
Contenu	Décalage concept de l'utilisateur / données monde réel	Non adapté au contexte Long temps de traitement Difficultés à extraire les informations importantes
Qualité	Par rapport au monde réel Par rapport aux tâches	Temps de correction Perte de confiance
Structure & Interopérabilité-	Modélisation des trajectoires, *etc.*	Long temps de traitement Mauvaise précision
Format fichier	Par exemple : différents systèmes de référence	Traductions / transformations nécessaires
Volume		Frustration, ralentissement
Présentation	Symbologie, sémiologie, texte, *etc.*	Interprétation compliquée
Coût & Licence		Non utilisation
Diffusion	Par exemple : numérique, papier, *etc.*	Mises à jour nécessaires Long temps de traitement
Sélectabilité	Réduire les données utilisées selon l'utilisation	Long temps de traitement

Afin d'améliorer l'usabilité de l'information, Brown et al. (Brown et al. 2013) proposent de faire évoluer deux aspects complémentaires : d'un côté améliorer le design de l'information elle-même et ses caractéristiques ; de l'autre côté, la façon de l'utiliser (visualisation et manipulation). Ces travaux ont porté sur l'utilisation de l'information géographique, mais ils sont tout aussi valables pour tout autre type d'information.

Dans cette seconde partie du chapitre, nous avons présenté les travaux relatifs à l'acceptation de la technologie et des données, notamment dans le cas de l'utilisation d'information géographique pour la prise de décision. Ces travaux identifient particulièrement les limites cognitives des utilisateurs, leurs habitudes et le contexte socio-technique dans lequel ils évoluent comme possibles freins à cette utilisation de nouvelles technologies. Maintenant que nous avons identifié ces limites sur l'usage et l'usabilité de l'information et de la technologie, nous pouvons présenter les résultats de travaux qui s'intéressent au rôle et à l'utilisation de l'information géographique, en particulier. La prochaine partie de ce chapitre porte donc sur l'importance de l'information **géographique** pour la gestion des risques, ainsi que sur son rôle majeur dans l'analyse des évènements, la communication, la planification et la prise de décision en situation de crise (Brown et al. 2013).

1.3. Le rôle de l'information géographique pour la gestion des risques

1.3.1. Le processus de raisonnement par la carte

Cette partie présente les objets d'étude qui sont rendus accessibles par l'information géographique et la cartographie / visualisation d'information géographique. Nous menons ainsi un état de l'art des différents processus de raisonnement par la carte qui ont été proposés par la communauté de l'information géographique, afin de le comparer aux tâches de la gestion des risques.

Depuis des centaines d'années, la carte a eu pour but premier de représenter l'espace géographique, tel qu'avec les cartes topographiques et les cartes dressées par les grands explorateurs du globe. Par la suite, la carte est alors devenue un outil pour l'analyse et la représentation d'information thématiques, et non plus seulement sur le monde concret qui nous entoure (Antoni et al. 2004 ; Brunet 1987). Le rôle de la carte comme outil de représentation thématique était alors de mettre en évidence des « *phénomènes non directement visibles* » (Antoni et al. 2004 ; Béguin et Pumain 1994) et interpréter leurs structures spatiales (Pumain et Saint-Julien 1997). La carte comme objet « concret » a donc eu pour première utilisation la **communication d'informations** préalablement étudiée et assimilées (MacEachren 1995 ; Bertin 1967). Aujourd'hui, les très grandes quantités d'informations disponibles ont inversé le problème, et l'utilisation première de l'information et de sa représentation a pour finalité **l'analyse** et **exploration**, afin d'en extraire de nouvelles connaissances (MacEachren 1995 ; MacEachren et al. 1999 ; Kraak et Ormeling 1996). Suite à ces nouveaux buts d'utilisation, les Systèmes d'Information Géographique (SIG) ont alors pris le devant sur la carte « statique », en permettant une analyse et représentation des données géographiques au sein d'un même outil logiciel. Notre façon de voir et modéliser l'information géographique a donc évolué d'une simple structure topographique et géolocalisée, vers un concept complexe d'évènement spatio-temporel.

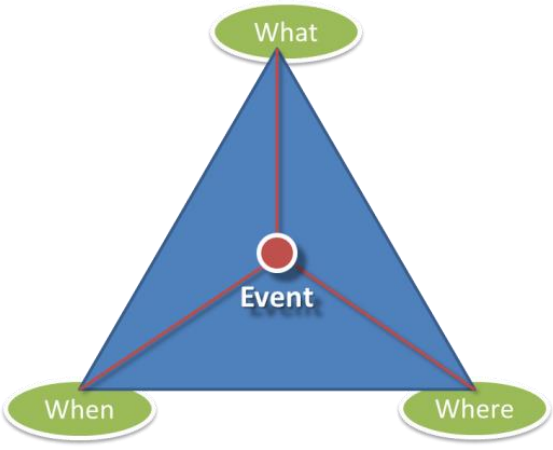


Figure 1-18. Modèle de la triade pour la définition d'un évènement (d'après Peuquet 1994)

Afin de définir un évènement spatio-temporel, Peuquet (Peuquet 1994) utilise la notion de « triade » (Figure 1-18), c'est-à-dire un concept défini par trois propriétés : Quoi, Où et Quand. Les composantes spatiales, temporelles et attributaires (sémantiques) sont ainsi prises en compte de manière liées pour décrire et expliquer un évènement. Pour l'analyse de ce type d'informations, Peuquet (Peuquet 1994) identifie alors trois grandes classes de questions qui peuvent servir de base

pour formuler des requêtes spatio-temporelles complexes. Ces classes sont approfondies dans l'article plus récent d'Andrienko, Andrienko et Gatalsky (Andrienko et al. 2003). Ces trois classes correspondent à des requêtes pouvant avoir l'une des formes suivantes :

- *Quand + Où → Quand* : recherche d'information sur les dates où certains évènements ont pu être détectés dans une zone donnée.
- *Quand + Quoi → Où* : recherche d'information sur les lieus où certains évènements ont pu être détectés dans une période de temps donnée.
- *Où + Quoi → Quand* : recherche d'information sur les objets et évènements qui ont pu être détectés dans une zone donnée et dans une période donnée.

Au sein de ces requêtes, les niveaux d'abstraction de ces trois axes peuvent être différents, comme un forte précision sur le temps et une large marge sur l'espace (Koussoulakou et Kraak 1992 ; Andrienko et al. 2003), que l'on retrouverait dans une requête du type : c*omment ont évalué les valeurs sur l'entité géographiques E ?*

De cette manière, l'auteur compare les requêtes rendues possibles par cette triade de données, par rapport aux buts d'utilisation des systèmes d'information géographique (SIG) tels qu'ils ont été conçus à leurs débuts : **exploration** de données, **explication / interprétation** (analyse), **prévision** et **préparation** (décision). Ces étapes sont aussi identifiées par Kimerling et al. (Kimerling et al. 2009) dans leur définition des étapes d'utilisation de la carte, dans le sens général de cet objet : **lecture**, **analyse** et **interprétation**. La première étape correspond à l'assimilation des symboles et du phénomène cartographié ; l'analyse correspond aux lectures et mesures effectuées sur la carte ; l'interprétation cherche à donner du sens à ces phénomènes observés. Les étapes d'utilisation de l'information géographique et de la carte identifiées par Peuquet (Peuquet 1994) et Kimerling et al. (Kimerling et al. 2009) rappellent fortement les étapes de gestion de l'information pour la prise de décision, telle que nous avons pu l'expliquer plusieurs fois dans ce chapitre. Kimerling et al. reviennent sur l'importance de prendre en compte plusieurs dimensions en compte (comme la « triade » de Peuquet) et différents types de connaissances afin d'interpréter les données géographiques :

> *It is impossible to interpret [maps] in isolation. (Kimerling et al. 2009, xxvii)*

L'analyse spatiale propose un ensemble d'outils, des techniques et de visualisations qui permettent ainsi de mieux comprendre la répartition des objets dans l'espace et le temps : de cette manière sont identifiées les grandes structures spatiales caractérisées par les continuités et les ruptures, les variations continues, et les oppositions majeures (Pumain et Saint-Julien 1997). Ces études dépendent bien sûr de l'échelle des phénomènes étudiés et du niveau de détail de ces informations : le problème MAUP (*Modifiable Areal Unit Problem*) en est une conséquence notable (Gehlke et Biehl 1934 ; Openshaw 1983). Leclerc (Leclerc 2002) montrent l'effet du MAUP sur l'étude de données de pauvreté, ainsi que le dilemme posé entre précision géographique et précision des statistiques. La Figure 1-19 illustre le problème MAUP, par la cartographie d'un même indicateur mais en utilisant des maillages différents : les mêmes données sont agrégées et représentées au département ou à la commune, au Honduras. Nous pouvons y observer des résultats largement différents selon le découpage géographique.

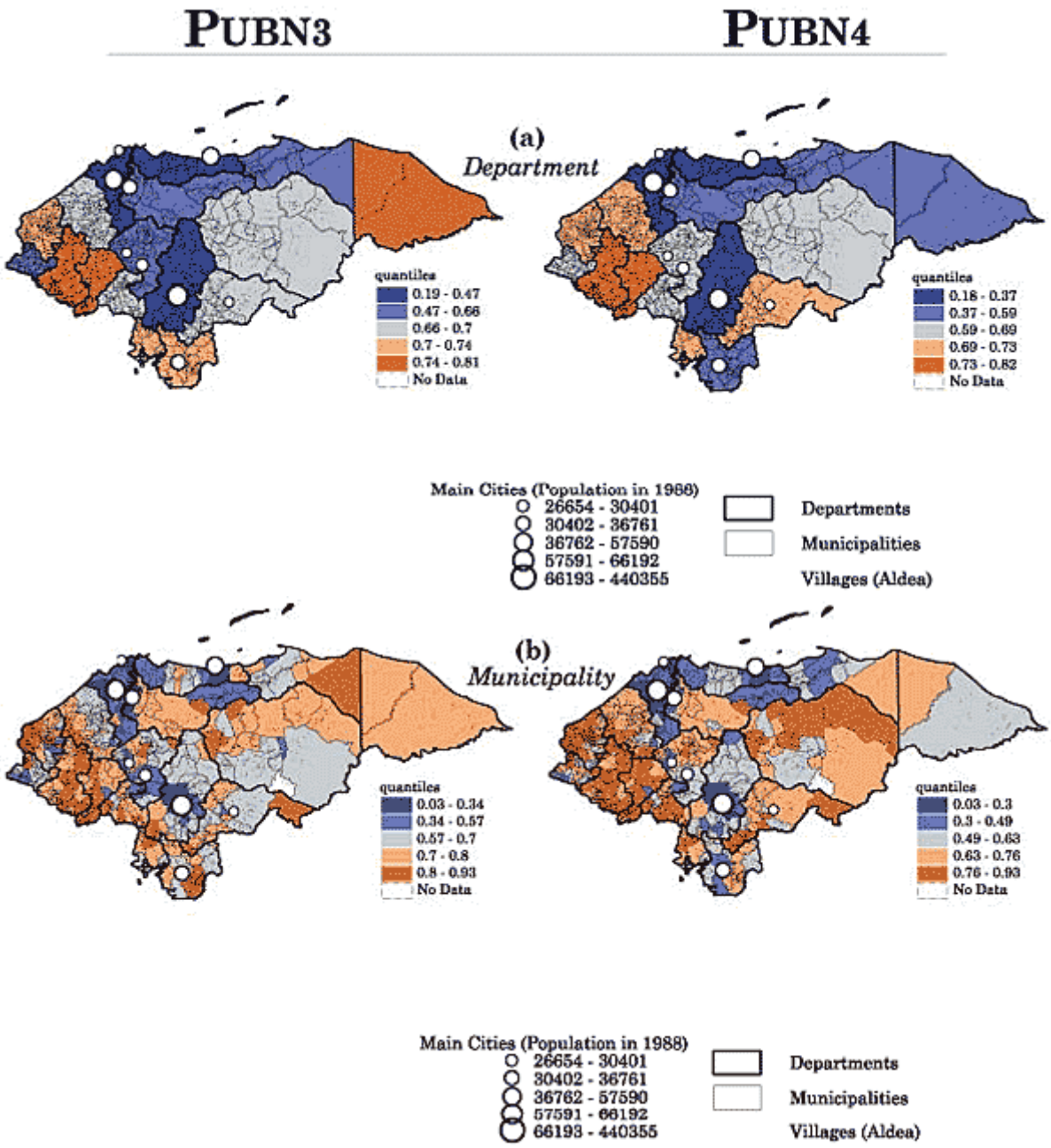


Figure 1-19. Illustration du MAUP pour l'étude de la pauvreté au Honduras : deux indices à l'échelle du département (a) ou de la commune (b) (*source :* Leclerc 2002)

Cette partie nous a permis d'identifier le concept d'évènement spatio-temporel, basé sur des informations géographiques, qui est au cœur du processus de raisonnement par la visualisation. Afin d'avoir la meilleure compréhension possible des structures de ces évènements, l'analyse visuelle doit à la fois permettre d'identifier les structures dans l'espace, le temps, ainsi que dans les attributs des objets étudiés. Dans la prochaine partie, nous nous intéressons aux différentes stratégies pour étudier les mouvements d'objets mobiles, dans l'espace géographique et dans le temps.

1.3.2. L'analyse des données de mouvement par la visualisation

1.3.2.1. De l'exploration des mouvements à l'analyse du contexte

Les nombreux travaux menés par Andrienko et Andrienko ont porté sur l'analyse des mouvements d'objets mobiles (trajectoires), leur distribution dans l'espace et leur caractérisation par rapport au contexte. Pour cela, plusieurs stratégies ont été proposées pour l'analyse spatiale de données de mouvement, transformant ainsi les données d'origine en connaissances. Le processus de base pour l'analyse de trajectoires s'intéresse avant tout au passage d'informations de mouvement vers des informations d'évènements, et enfin des informations de zones caractéristiques (Andrienko et al. 2011). La Figure 1-20 présente le processus d'analyse spatiale proposé par ces auteurs, depuis les données de base jusqu'à l'identification de groupes logiques dans le mouvement. Ce processus se compose des étapes suivantes :

- **Extraction d'évènements depuis les données de mouvement :** les données de mouvement de base ne correspondent pas à de l'information d'intérêt pour l'étude du mouvement. Les évènements tels que les lieux d'arrêt, les moments d'accélération, etc., permettent de décomposer les trajectoires selon des motifs d'intérêt (Laube et al. 2005a ; Dodge et al. 2009 ; Vandecasteele et al. 2014).
- **Regroupement spatial des évènements, reconnaissance des lieux d'intérêt :** à partir de l'étude la densité des positions d'intérêt, des zones sont identifiées comme celles à forte fréquentation, ou faible fréquentation (Willems et al. 2010 ; Idiri et Napoli 2012b ; Andrienko et Andrienko 2011 ; Demšar et Virrantaus 2010)
- **Agrégation des évènements et données de mouvement :** les données de mouvement peuvent être regroupées selon plusieurs caractéristiques communes, comme leur direction, leur vitesse, le type d'objets concernés, etc. Un exemple cartographique de base serait une carte de flux, basée sur un découpage de l'espace en zones (Andrienko et Andrienko 2013 ; Andrienko et Andrienko 2010 ; Andrienko et Andrienko 2011 ; Scheepens et al. 2014).
- **Analyse de l'agrégation des données de mouvement :** des méthodes beaucoup plus interactives doivent être utilisées au cours de cette étape d'interprétation, en permettant une exploration complète des informations sous toutes leurs dimensions (Hurter et al. 2009 ; Guo et al. 2011 ; Riveiro 2011). Cette étape est donc nécessaire afin d'interpréter les regroupements de données identifiés au cours des étapes précédentes.

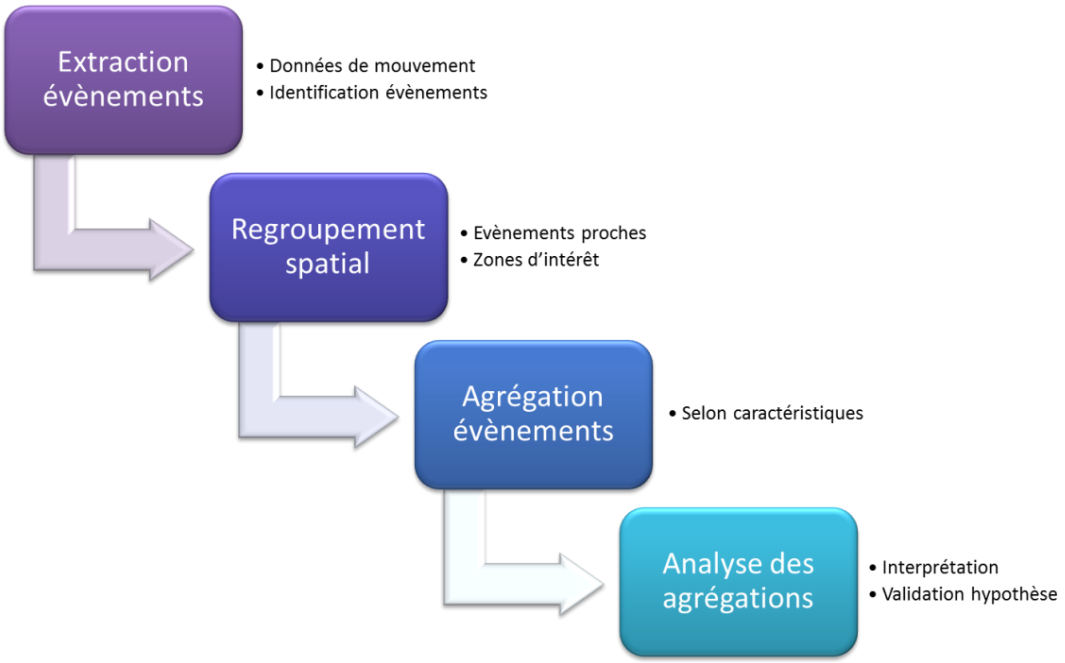


Figure 1-20. Processus d'analyse spatiale de données de mouvement (d'après Andrienko et al. 2011)

En 2013, les auteurs proposent un nouvel état de l'art des techniques de visualisation d'information de mouvement, en prenant en compte le contexte des trajectoires (Andrienko et Andrienko 2013). Dans cet article, de nouvelles étapes sont proposées pour l'analyse de ce type de données, identifiant quatre étapes successives à différents niveaux de détails (Figure 1-21).

- **Visualisation des trajectoires :** visualisation des trajectoires dans leur ensemble, regroupement général des mouvements. Les trajectoires ne sont pas regroupées, mais toutes considérées de manière exhaustive (cartographie de l'ensemble des trajectoires).
- **Exploration au sein des trajectoires :** approfondir l'exploration des attributs, des motifs et des évènements identifiés dans l'espace étudié. Les trajectoires sont découpées de manière sémantique : cette étape est proche de la deuxième dans le processus précédent.
- **Vue d'ensemble sur le mouvement :** analyser la densité, la présence des objets, généraliser les mouvements à l'échelle des groupes, en prenant en compte le temps et l'espace.
- **Analyse des mouvements selon le contexte :** comparaison des informations de mouvement par rapport au contexte où elles se situent. Ce contexte doit permettre de donner du sens à l'information, tel que le contexte géographique, météorologique, etc. Cette étape vient spécifier la quatrième étape du processus précédent, en s'intéressant aux trajectoires une par une.

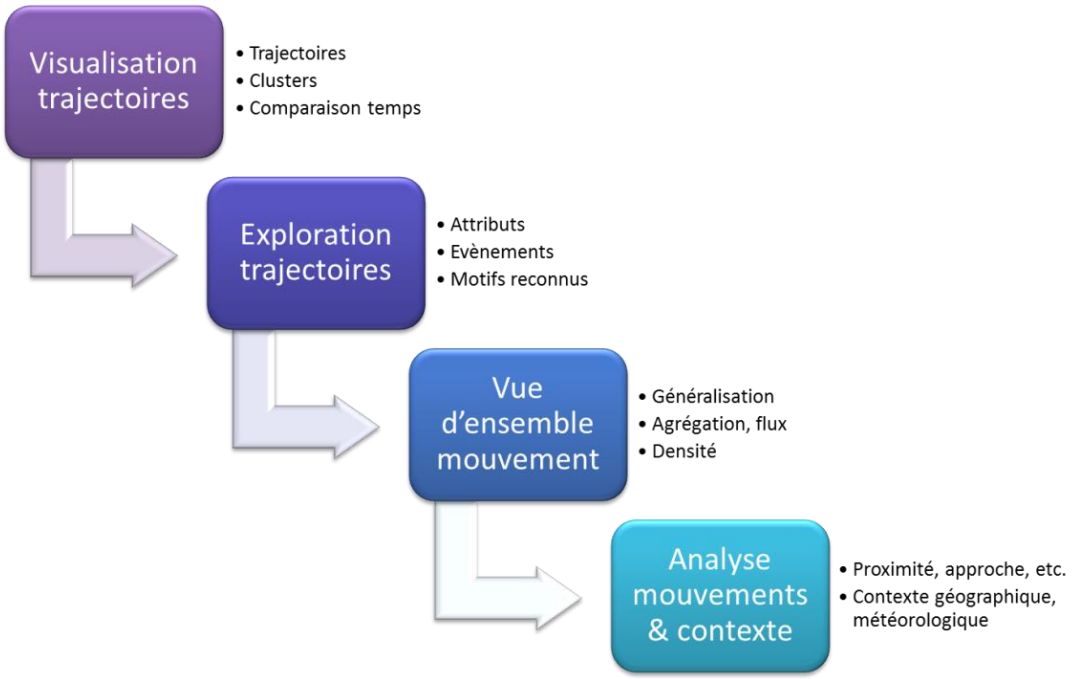


Figure 1-21. Processus d'analyse de trajectoires (d'après Andrienko et Andrienko 2013)

Tout comme les étapes proposées par Kimerling et al. (Kimerling et al. 2009), ces quatre étapes proposées pour la visualisation et l'exploration des données spatio-temporelles se font vers un niveau de détail de plus en plus fin, ainsi qu'un nombre de données de plus en plus restreint. En effet, la répartition générale des mouvements est déjà visualisée et analysée, sans tri dans les informations de base. L'exploration s'intéresse ensuite plus particulièrement aux relations entre l'espace, le temps, puis enfin les interactions entre les comportements des objets et le contexte dans lequel ils évoluent.

Dans le domaine de la surveillance maritime, Davenport et Risley (Davenport et Risley 2006) définissent la notion de MISR (*Maritime Intelligence, Surveillance & Reconnaissance*) de manière à rassembler l'ensemble des outils et technique pour l'analyse des mouvements des navires en mer, pour la défense nationale. Au sein de la MISR, les auteurs identifient de nombreux axes de recherche fondamentaux pour l'amélioration de l'étude des mouvements de ces objets en mer, tel qu'il est illustré sur la Figure 1-22. A partir des limites posées par les systèmes de surveillance actuels, Davenport et Risley mettent en évidence quatre axes de recherche en visualisation pour des applications à la défense, que sont la **visualisation d'information spatio-temporelle**, la **visualisation d'information connue**, la **visualisation de ce qui est inconnu**, et enfin la **qualité et l'incertitude** liées à ces informations. Ces domaines de recherche permettent, entre autres, de mieux comprendre les comportements des navires en mouvements, détecter les comportements anormaux, émettre des alertes lors de l'approche de zones sensibles ainsi que les données essentielles des navires surveillés.

Toutes ces étapes et informations sont nécessaires à la prise de décision, relative à la sécurité et/ou sûreté maritime. Les auteurs s'intéressent principalement à la cartographie et la visualisation de ces informations, afin de supporter les opérateurs dans leurs tâches quotidiennes. Les axes d'automatisation pour la reconnaissances de comportement sont notamment soulevés dans leur travaux avec la collaboration de Roy (Roy 2008 ; Roy 2009 ; Roy et Davenport 2009).

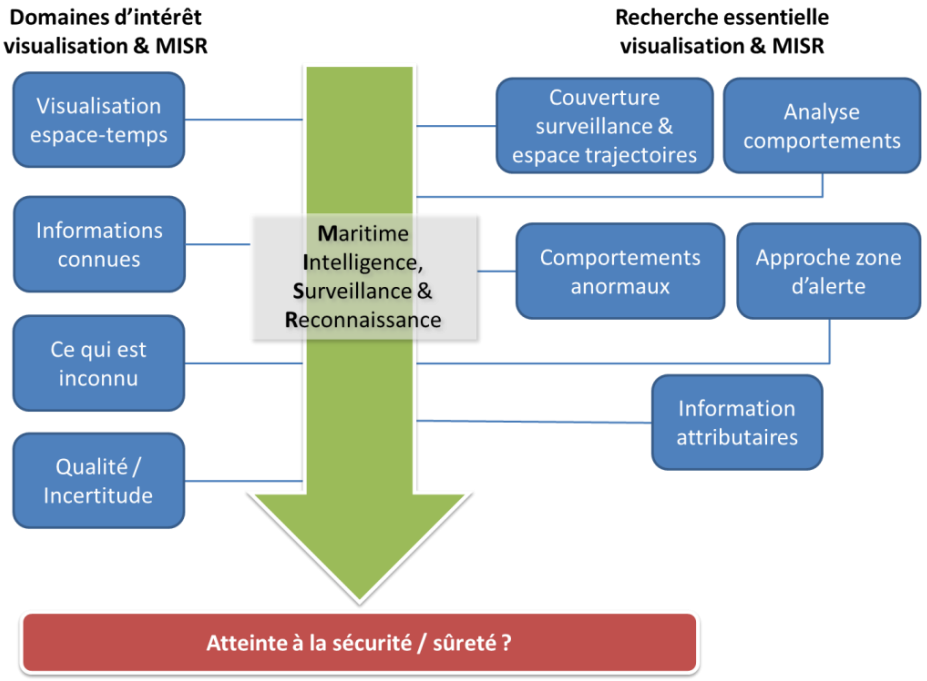


Figure 1-22. Définition de la MISR (d'après Davenport et Risley 2006)

Ces nombreux travaux présentés ici, allant de la cartographie générale à l'étude des comportements de navires, montre la grande importance de la visualisation de la dynamique et des comportements des navires, pour la prise de décision et la gestion des risques en mer. Ces étapes peuvent facilement être extrapolées à des domaines différents mais proches, tels que la gestion du trafic aérien, qui évolue dans un espace en 3 dimensions, ou la gestion du trafic routier, qui dépend d'un réseau défini (les routes « physiques »).

1.3.2.2. Les comportements à risques, objets d'étude cartographique

Ces nombreux processus d'analyse de l'information géographique pour l'étude de mouvements font souvent appel à la notion d'évènement, dont nous avons expliqué la définition grâce aux travaux de Peuquet (Peuquet 1994). De nombreux auteurs, spécialisés dans la fouille de données, ont pu identifier plusieurs motifs récurrents dans l'analyse de trajectoires, qui vont pouvoir être identifiés par rapport au reste des données (Laube et al. 2005b ; Gudmundsson et al. 2012). Ces motifs retrouvés dans les trajectoires se basent sur la direction et la localisation des entités mobiles, et peuvent être classés en cinq groupes distincts : la convergence, la rencontre, les flux communs, la localisation fréquente (récurrence) et le *leadership* (voir Figure 1-23). Les définitions de ces termes, donnés par Laube et al. (Laube et al. 2005b), sont ensuite modifiées et précisées par Benkert et al. (Benkert et al. 2008). Ces motifs peuvent ensuite être complétés par d'autres notions mises en avant dans la littérature : par exemple, les notions d'anomalie (*outlier*) et de motif périodique (Benkert et al. 2008 ; Etienne et al. 2012 ; Idiri et Napoli 2012a).

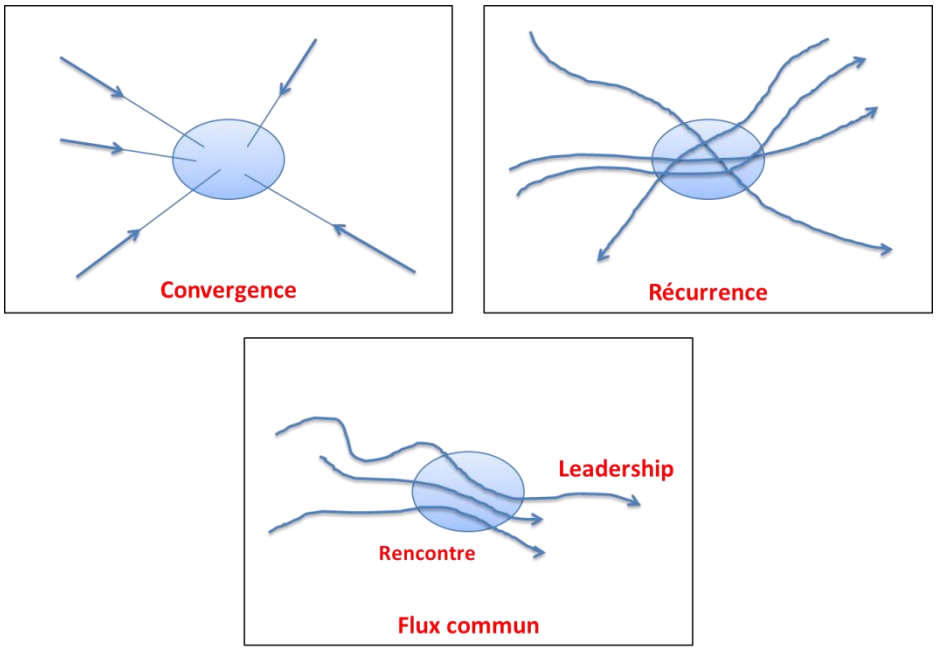


Figure 1-23. Différents motifs récurrents en analyse de trajectoires (d'après Bogorny 2010)

En parallèle à la notion de motifs spatio-temporels, la notion de comportement d'objets mobiles doit être éclaircie. Ce que nous appelons « comportement » est la dynamique d'un certain objet mobile au sein d'un contexte donné : dans le cadre de l'étude des navires, Idiri (Idiri 2013, 34) définit un comportement comme *l'ensemble des mouvements d'un objet au sein d'un contexte géographique donné*. Dans cette définition, la *situation* est une combinaison de caractéristiques de l'objet étudié, de son contexte et de son environnement d'évolution spatio-temporelle. Un comportement peut donc être un changement de vitesse, un changement d'attribut, ou encore un motif défini clairement et retrouvé au sein d'un ensemble d'objets mobiles. Par exemple, nous pouvons dire qu'un navire qui se rapproche d'un autre navire va adopter un comportement d'intérêt, qui demande une analyse afin d'en comprendre les raisons et conséquences ; tout comme un navire qui sort d'une route maritime. L'interprétation de ce comportement nécessite donc d'analyser les propriétés géométriques, temporelles et sémantiques de sa trajectoire, au sein de son environnement d'évolution (Yan et al. 2008 ; Parent et al. 2013 ; Vandecasteele et al. 2014). Afin de comprendre le comportement d'un objet, de nombreuses informations sur son déplacement et son contexte d'évolution doivent donc être croisées afin de reconnaitre ou non un certain motif d'intérêt (Idiri et Napoli 2012a ; Idiri 2013).

En définissant le MISR tel que nous l'avons précédemment, Davenport et Risley (Davenport et Risley 2006) montrent la nécessité de visualiser le comportement des navires, ce qu'ils définissent par la position et les mouvements passés des objets mobiles, dans le temps et l'espace. Un évènement d'intérêt que les auteurs qualifient d'inexistant dans les systèmes de surveillance actuel est, par exemple, le croisement de deux trajectoires dans le temps et l'espace (Davenport et Risley 2006, 100), ou bien des trajectoires proches et parallèles (Davenport et Risley 2006, 105). Ainsi, les auteurs mettent en avant les besoins de visualiser par la carte les comportements dits « normaux » des navires, afin de mieux retrouver et comprendre les comportements « anormaux » et possiblement à risques (Lavigne et al. 2011).

Conclusion du Chapitre 1

Dans ce chapitre, nous avons présenté de manière générale la notion de risques au sein de systèmes, et la façon dont l'être humain est amené à étudier, analyser et gérer ces évènements porteurs de danger. Nous avons orienté notre étude vers la contribution de la technologie lors de l'analyse des comportements à risques, et lors la prise de décision. Pour cela, nous avons présenté le rôle des systèmes d'aide à la décision (SAD) et les environnements d'aide à la résolution de problèmes (PSE). Les SAD sont des systèmes d'information qui supportent l'analyse et la prise de décision dans des situations complexes. Ils sont fondés sur une formalisation des connaissances du domaine, afin de proposer une expertise sur une classe de problèmes donnés, à partir de faits et règles. Afin de prendre en compte de nombreuses possibilités, les SAD sont, par définition, des outils flexibles et adaptables selon le contexte d'utilisation. Un SAD repose donc principalement sur le modèle de l'environnement ciblé, un moteur de simulation et une interface de présentation des informations gérée par un moteur de recommandation. Les PSE se basent sur des méthodes de calcul automatisées, afin de guider dans la résolution de problèmes identifiés (Gallopoulos et al. 1992 ; Gallopoulos et al. 1994). Pour cela, les PSE utilisent des outils de simulation tout en conservant l'historique des actions menées, afin d'améliorer la simulation à venir. Les PSE sont notamment utilisés afin de résoudre des problèmes de manière rapide et de mener des actions originales, non identifiées précédemment.

Nous avons aussi étudié les nombreuses stratégies existantes pour la gestion des risques, qui peuvent prendre en compte la dimension spatiale dans leur stratégie. Nous avons notamment présenté, au sein d'autres méthodes qui supportent la cognition de l'utilisateur, les systèmes cognitifs joints (JCS). Ceux-ci mettent en avant la nécessité d'utiliser à la fois la puissance des machines pour le calcul et la simulation, couplée au raisonnement et aux connaissances de l'être humain. En revenant sur ce principe, nous pouvons qualifier les SAD et les PSE de systèmes cognitifs joints, grâce aux interfaces et calculs pour améliorer la prise de décision par l'homme.

A la fin de ce chapitre, nous avons ouvert sur l'utilisation de l'information géographique pour la prise de décision, une source d'information qui se trouve être fondamentale pour l'étude et la prévention des risques liés aux comportements d'objets mobiles. L'analyse et la compréhension des mouvements d'objets doit en effet permettre de comprendre la différence entre des comportements normaux, à définir, et des comportements différents, potentiellement à risques. Pour cela, nous avons orienté cette recherche vers le rôle important de l'analyse de cette information spatiale, tout en autorisant l'être humain à garder le contrôle de la situation et des décisions effectuées. En effet, tout comme le préconisent les JCS, la décision la plus efficace est celle qui peut être effectuée par la collaboration entre la machine, grâce à sa puissance de calcul, et l'être humain, grâce à ses connaissances dans le domaine étudié.

Toutefois, les SAD et PSE « classiques » sont limités dans leur utilisation de la visualisation d'information, notamment à caractère géographique. En effet, l'analyse géovisuelle est un domaine de recherche très récent, dont les nombreuses possibilités n'ont pas encore été prises en compte dans de tels systèmes opérationnels pour l'analyse des risques. Cela se retrouve notamment dans les projets les plus récents en gestion de risques : l'application au domaine de la surveillance maritime, par exemple, favorise essentiellement les capacités à raisonner automatiquement sur de grandes

quantités de données, pour mieux repérer les situations à risques. La visualisation d'information et le rôle de l'être humain se retrouvent alors mis au second plan dans ces études. Dans le prochain chapitre, nous revenons plus en détail sur les possibilités d'analyser et comprendre l'information spatiale, temporelle et sémantique, grâce du domaine de l'analyse géovisuelle appliqué à l'étude de mouvements d'objets.

CHAPITRE 2 - ANALYSE GEOVISUELLE ET FORMALISATION DE LA VISUALISATION

Au cours du premier chapitre, nous avons présenté les différents aspects de la gestion des risques par l'analyse, la détection, la prise de décision et la prévention. Plusieurs approches et modèles de ce processus ont été comparées, afin d'identifier les étapes fondamentales de l'analyse des évènements qui peuvent mener au risque. A la suite de cela, le rôle de l'information géographique dans la prise de décision au sein de la boucle décisionnelle a été mis en évidence.

Dans ce chapitre, nous présentons le domaine de l'analyse géovisuelle, ainsi que ses nombreux domaines d'application. Nous nous intéressons particulièrement à l'exploration et à l'analyse des mouvements d'objets mobiles et des trajectoires qui leur sont associées. L'objectif de ce chapitre est de mener un état de l'art de la théorie de la visualisation, ainsi que la modélisation des techniques de visualisation d'information spatiale, temporelle et sémantique, pour l'étude de données de mouvement.

Dans une première partie, nous présentons la théorie de la visualisation d'information, pour ensuite orienter vers les sciences relatives à l'analyse visuelle d'information géographique. La deuxième partie de ce chapitre présentera les travaux menés ces dernières décennies afin de modéliser ces méthodes de visualisation, afin d'assister les usagers dans leur utilisation pour l'étude de données abstraites, scientifiques ou géographiques. La troisième et dernière partie présente la notion de formalisation de connaissances par des ontologies, un concept qui sera ensuite exploité pour la modélisation des environnements d'analyse géovisuelle.

2.1. De la représentation de données à l'analyse géovisuelle

Avant d'introduire la notion de représentation d'information géographique, il est important de savoir quels sont les modes de fonctionnement de notre perception, et de notre compréhension de l'information visuelle. Grâce à cette partie théorique, nous pouvons ainsi identifier quels sont les aspects clés sur lesquels travailler, afin de représenter des informations telles que les déplacements d'objets mobiles et ainsi améliorer le travail de surveillance et de contrôle de ces informations.

2.1.1. Perception de l'information

L'étude du fonctionnement de la perception et de la vision humaine est une étape importante afin de définir le rôle et la compréhension de la visualisation d'information. Il est donc nécessaire d'étudier la façon dont nous percevons les informations visuelles, avant de pouvoir développer de nouvelles méthodes de représentation basées sur la perception visuelle. Les théories fondamentales qui permettent d'étudier l'organisation et la compréhension d'information visuelle remontent au début du XX^{ème} siècle.

La perception visuelle peut notamment être définie comme la **reconnaissance des formes** représentées et de leur organisation structurale. Ce concept se situe donc au niveau cognitif de la vision : c'est cette étape qui permet de donner un sens aux formes et aux objets (stimuli visuels) que notre système visuel **perçoit**, et que notre esprit **reconnait** (Card et al. 1999).

Nous présentons ici trois des théories les plus importantes en perception visuelle, qui sont utilisées lors de la représentation d'informations visuelles : la perception pré-attentive, les lois de la Gestalt et la théorie de la sémiologie graphique de Bertin.

2.1.1.1. La perception pré-attentive

Deux types de perception sont identifiés : la **perception pré-attentive** ou automatique, qui prend place en amont de la **perception attentive**, qui elle est plus tardive et contrôlée. La perception pré-attentive se base sur notre mémoire à court terme et sur un traitement mental très rapide et sans effort, grâce aux attributs visuels les plus simples des objets visualisés (Treisman et Gormican 1988 ; Few 2004). Les résultats de cette perception sont donc instantanés. La Figure 2-1 montre la place de la mémoire à court terme dans le processus de perception visuelle et de la pré-attention de notre système visuel.

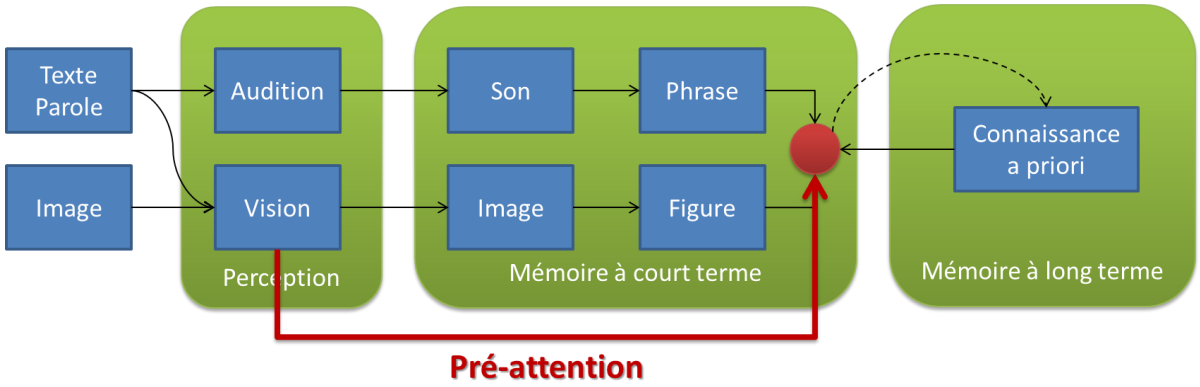


Figure 2-1. Fonctionnement de la perception pré-attentive

La perception pré-attentive utilise régulièrement un attribut visuel qui est unique et distinct du reste des objets visualisés (Healey et al. 1996). Par exemple, la Figure 2-2 (a) met en évidence un objet d'une couleur différente que nous reconnaissons automatiquement comme distinct. Sur la Figure 2-2 (b), il est plus compliqué de retrouver ce même cercle noir car la variable visuelle de la couleur noire est utilisée par de nombreux autres objets.

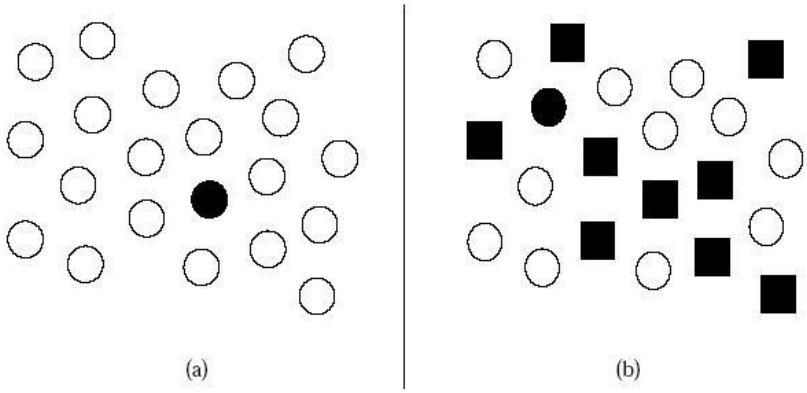


Figure 2-2. Détection d'un objet distinct par la perception pré-attentive (*source :* Healey et al. 1996)

Afin d'aider le travail de la perception pré-attentive, plusieurs théories ont été identifiée par la communauté de la psychologie cognitive, de la représentation d'information, et de la cartographie. Les sous-parties suivantes présentent donc les lois et concepts qui permettent de soutenir une telle perception automatique de l'organisation de l'information visuelle.

2.1.1.2. Les lois de la Gestalt : psychologie de la forme et regroupements

La reconnaissance d'objets et d'organisation dans un ensemble d'objets suit aussi des lois connues comme la « psychologie de la forme », ou lois de la Gestalt depuis les années 1920. Cette théorie est une synthèse mentale qui favorise la perception du « tout » plutôt que la « somme des parties » (Köhler 1964). Dans ces lois, on retrouve la notion des niveaux de lecture d'une scène, tout comme dans le monde de la cartographie : un niveau global dans la lecture général d'un groupe d'objet, ainsi qu'un niveau plus local où un objet va être lu et interprété de manière isolée par rapport au reste du groupe d'objets représentés.

La Gestalt est constituée de plusieurs lois de perception, qui peuvent être utilisées pour simplifier le regroupement visuel d'objets et ainsi améliorer leur distinction : la *loi de la bonne forme*, la *loi de continuité*, la *loi de proximité*, la *loi de similitude*, la *loi de destin commun* et la *loi de familiarité*. Il n'y a pas de loi prépondérante parmi celles-ci, elles peuvent donc agir simultanément. Parfois, cela peut mener à une contradiction quant à l'interprétation que nous avons des ensembles qui ressortent de l'organisation des objets (Metz-Göckel 2012). La Figure 2-3 illustre un cas de contradiction d'interprétation : notre esprit distingue deux formes connues, mais ne peut pas déterminer une unique interprétation des objets présentés.

La contradiction est due à l'utilisation de formes connues (visages noirs, vase blancs), qui sont alors automatiquement interprétées comme telles. La loi de la forme est en effet la plus souvent utilisée en psychologie de la Gestalt, la plupart des lois suivantes étant des variantes de cette première loi. Elle stipule qu'*un ensemble de parties informe (des groupements aléatoires) tend à être perçu d'abord comme une forme, si cette forme se veut simple, symétrique, stable*.

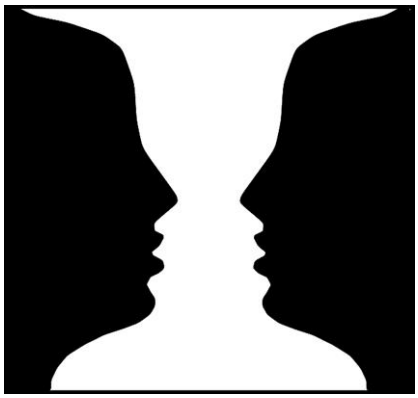


Figure 2-3. Vase de Rubin : contradiction sur l'interprétation des formes

Cela fait aussi appel à la *loi de familiarité*, selon laquelle nous percevons les formes qui nous sont les plus connues. Cette loi souligne l'importance de prendre en compte, lors de la visualisation d'information, la **familiarité** ou **connaissance** du domaine que l'on peut avoir : nous sommes plus aptes à reconnaitre ce que nous **connaissons déjà**. Un contre-exemple anecdotique de cette loi se retrouve dans l'histoire de Christophe Colomb : la légende dit que les indiens d'Amérique n'ont pas vu venir les frégates, car ils ignoraient l'existence de tels objets et ne pouvaient ainsi pas projeter cette réalité sur les formes qu'ils voyaient.

La Figure 2-4 représente la *loi de proximité* de cette théorie : intuitivement, nous regroupons les objets qui sont proches et nous séparons les objets éloignés. La figure présentée ici utilise une unique forme, mais en utilisant plusieurs symboles (cercles, carrés, triangles) nous pouvons observer que la loi de la proximité prévaut sur la similitude de variables visuelles telles que la forme ou la couleur. La Figure 2-5 illustre la *loi de clôture* : l'esprit peut détecter automatiquement une forme qui n'existe pas, mais qui est simplement suggérée par une partie de son contour. Ainsi, avant même d'utiliser les variables visuelles classiques, la théorie de la Gestalt permet de faire des rapprochements simples entre éléments, sans effort cognitif.

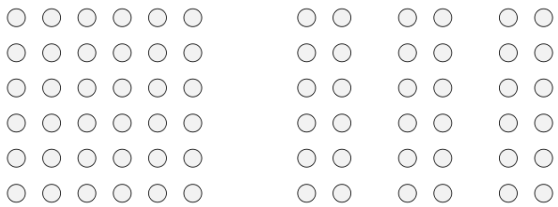


Figure 2-4. Loi de proximité

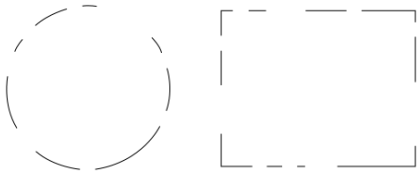


Figure 2-5. Loi de clôture

En utilisant ces deux grands principes, la visualisation d'information peut alors être facilitée : moins d'efforts devront être fournis pour comprendre l'information, elle sera ainsi rapidement acquise et mieux mémorisée. Ainsi, grâce à la théorie de la Gestalt, nous avons identifié certaines lois qui régissent notre perception pré-attentive de l'information visuelle, c'est-à-dire la façon dont nous comprenons automatiquement la présentation de groupements d'objets, d'après leur forme et leur disposition. Dans la sous-partie suivante, nous nous intéressons plus particulièrement aux différentes

variables visuelles qui permettent d'identifier des objets et leurs propriétés, qu'ils soient pris de manière isolée ou dans un groupe.

2.1.1.3. La sémiologie graphique : une utilisation efficace des variables visuelles

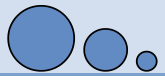
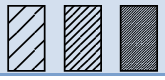
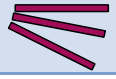
Les travaux fondamentaux sur les règles de représentation visuelle remontent à la formalisation de la sémiologie graphique proposée par Jacques Bertin (Bertin 1967), cartographe français. Cette sémiologie est résumée dans la Table 2-1. A l'époque, Bertin s'intéressait particulièrement à la représentation d'objets sur des cartes statiques (papier). Ses travaux d'origine ont permis d'identifier sept variables visuelles de bas niveau pour la représentation visuelle d'information, tout en conservant des pouvoirs de discrimination et d'association.

De façon à spécifier l'utilisation de ces variables visuelles, Bertin propose une classification pour le rôle de la variable visuelle, appelé « niveau », selon trois groupes :

- **Variable visuelle nominale (N) :** il s'agit d'une variable qualitative, qui permet de qualifier la place d'un objet spécifique dans un ensemble d'objets. La variable est dite sélective si elle permet de distinguer rapidement un objet parmi d'autres, ou *associative* si elle permet de regrouper certains objets par des attributs communs.
- **Variable visuelle ordinale (O) :** un tel type de variable permet de classer les objets selon une échelle de valeurs, selon des principes communs et intuitifs pour le public visé. De cette façon, l'usager peut répondre à des questions du type « Qui est plus grand que qui ? ». Par exemple, la taille des objets permet de créer un ordre du plus petit au plus grand.
- **Variable visuelle quantitative (Q) :** une variable est dite quantitative si, en plus d'un ordre, la quantité qui distingue deux objets peut être jugée grâce à un rapport numérique, comme par exemple : deux fois plus que (...), trois fois moins que (...), etc.

Sur la Table 2-1, nous évaluons aussi l'efficacité des variables visuelles de Bertin selon les niveaux d'organisation présentés précédemment, d'après les estimations de MacEachren (MacEachren 1995) : l'évaluation va de faible (0) à très efficace (++). Dans ce tableau, nous pouvons apprécier la valeur de la variable *Position* : cela revient à traduire de manière formelle la théorie de la Gestalt, selon laquelle la première façon de comprendre l'information visuelle est de distinguer des groupes de données, grâce à leur distribution spatiale. Cette capacité à pouvoir distinguer plusieurs objets est aussi appelée le « pouvoir » de la variable visuelle, dont la force augmente avec le nombre maximum d'objets que l'on peut différencier.

Table 2-1. Variables visuelles et évaluation de la perception

Type / Variable visuelle	Nominale		Ordinale	Quantitative
	Associative	Sélective		
Position (X, Y)	++	++	++	++
Taille 	0	+	++	++
Valeur	0	+	+	+
Grain 	+	+	++	+
Couleur	++	++	+	+
Orientation 	++	0	+	0
Forme	++	0	0	0

De nombreux travaux de recherche ont proposé leur propre évaluation du pouvoir perceptif de ces variables visuelles, pour la représentation d'informations statiques (Mackinlay 1986 ; Wilkinson 1999). Ces travaux s'intéressent notamment à la manière de représenter l'information selon le type de données considérées.

La façon de représenter l'information par moyens visuels dépend donc de nombreuses règles, qui permettent de structurer cette présentation de la manière la plus compréhensible possible. De plus, comme nous l'avons vu, les lois qui régissent la perception humaine de l'information visuelle sont connues. Grâce à celles-ci, nous pouvons faciliter l'assimilation des structures de données représentées, grâce à la reconnaissance automatique de propriétés visuelles. Ces grandes lois de la perception visuelle sont grandement utilisées dans les domaines qui s'intéressent à la représentation de l'information par des moyens visuels. Parmi ces moyens, la carte est un objet témoin important de l'évolution de notre société, notamment de notre manière de manipuler et de représenter l'information visuelle.

Maintenant que nous avons présenté les principales théories de représentation de l'information pour une perception optimale, nous présentons le cas concret de la cartographie. Afin d'analyser les comportements d'objets, qui est le cœur de cette thèse, la carte est un élément fondamental qui permet de visualiser et comprendre les dynamiques spatiales d'objets mobiles. Dans la prochaine sous-partie, nous présentons l'évolution du domaine de la représentation cartographique, de l'utilisation de l'information géographique et des fins de cette utilisation.

2.1.2. La cartographie, moyen représenter l'information connue

Depuis plus de 2000 ans, la représentation et l'étude de données spatiales se fait via des cartes, souvent considérées comme un art à part entière. La cartographie hérite en effet de nombreux siècles de témoignages, de réalisation, et plus récemment, d'une théorie à part entière. Cette sous-partie présente l'évolution du rôle et techniques de la représentation cartographique et des manières d'utiliser de telles représentations.

2.1.2.1. De la communication à l'exploration et l'analyse

La représentation cartographique a constamment évolué selon le contexte historique et selon son support. Jusqu'à la fin du XX^{ème} siècle, la cartographie était essentiellement un produit fixe dont le support était du papier, du métal, ou autres matériaux fixes. Jusqu'à cette période, ces cartes étaient alors statiques et ne présentaient pas d'interaction.

Néanmoins, la quantité d'information représentée n'en était pas pour autant réduite. Les cartographes ont toujours fait preuve d'une grande créativité afin de représenter, de manière plus ou moins facile à comprendre, les messages et les informations d'intérêt de la carte. Si nous reprenons l'exemple des portulans, dont le rôle était de permettre la navigation en mer, ces cartes portaient à la fois des informations sur les ports, la forme de la côte relative au point de vue du navire et les informations nécessaires à la navigation. Celles-ci comprennent par exemple les lignes de vents et leur cap et les zones de dangers en mer (courants, hauts-fonds, accidents, etc.) (Hofmann et al. 2012).

Historiquement, la représentation cartographique était essentiellement un moyen de communication de messages. Ces messages pouvaient concerner le rapport d'une exploration (maritime, terrestre), la représentation d'une certaine croyance religieuse ou mythique, ou bien plus tard, les résultats d'une analyse scientifique de l'espace : par exemple, l'étude de l'épidémie de choléra à Londres par John Snow en 1854, qui peut être considérée comme les prémices de l'analyse spatiale (voir Figure 2-6).

Un tout autre type de carte historique, qui présente un nombre importants d'informations, relève de ce que nous appelons aujourd'hui le *storytelling*, ou mise en récit. Le *storytelling* permet de représenter des évènements passés, sur la base d'une structure narrative proche du récit. On retrouve cette technique dans plusieurs recherches actuelles, liées à la visualisation d'information géographique, souvent utilisées pour retracer les évènements marquant d'une crise ou d'un accident (Ho et Jern 2013 ; Lundblad 2013). Mais cette façon de présenter l'information remonte à plusieurs centaines d'années : la *Carte figurative des pertes successives en hommes de l'Armée française dans la campagne de Russie en 1812-1813*, publiée en 1869 par Charles Joseph Minard, en est un parfait exemple (voir Figure 2-7).

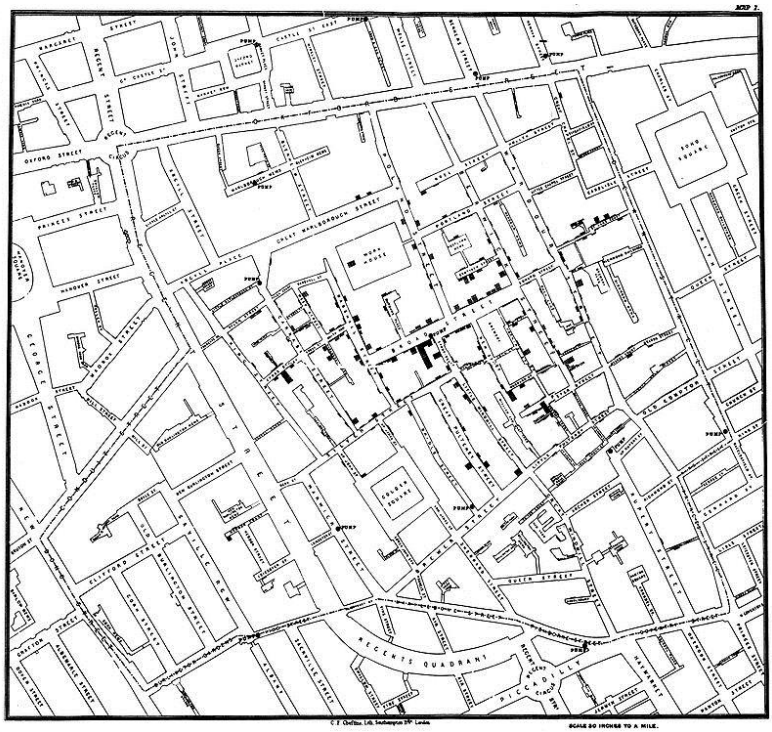


Figure 2-6. Amas de choléra pendant l'épidémie de choléra de Broad Street par John Snow (1854) (*source : Wikimedia*)

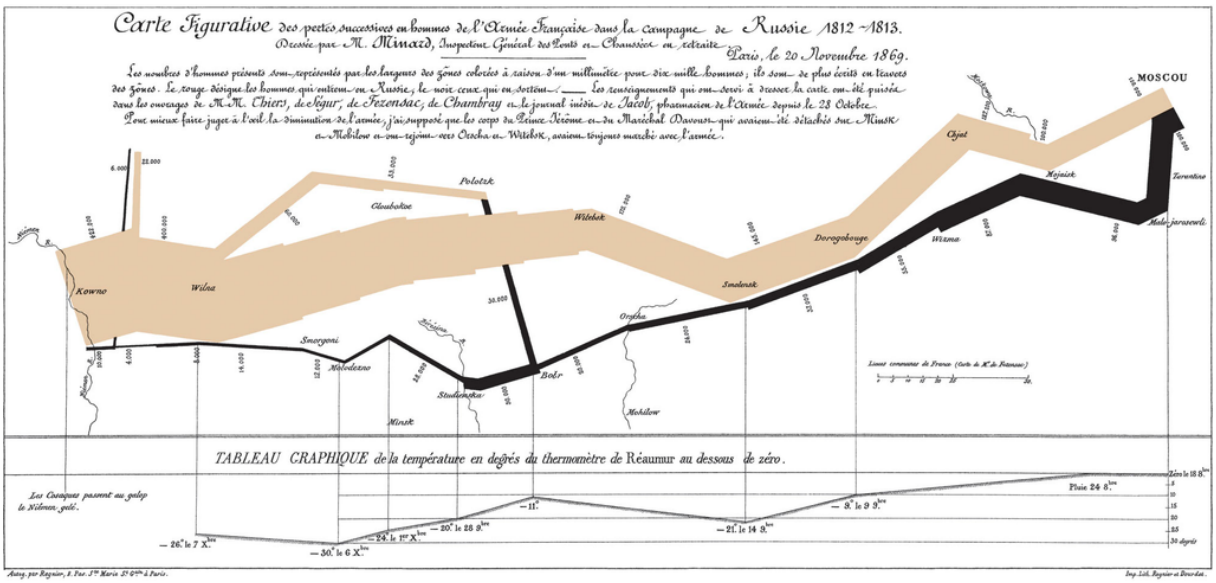


Figure 2-7. Carte de la campagne Napoléonienne en Russie, par Minard (1869) (*source : Wikimedia*)

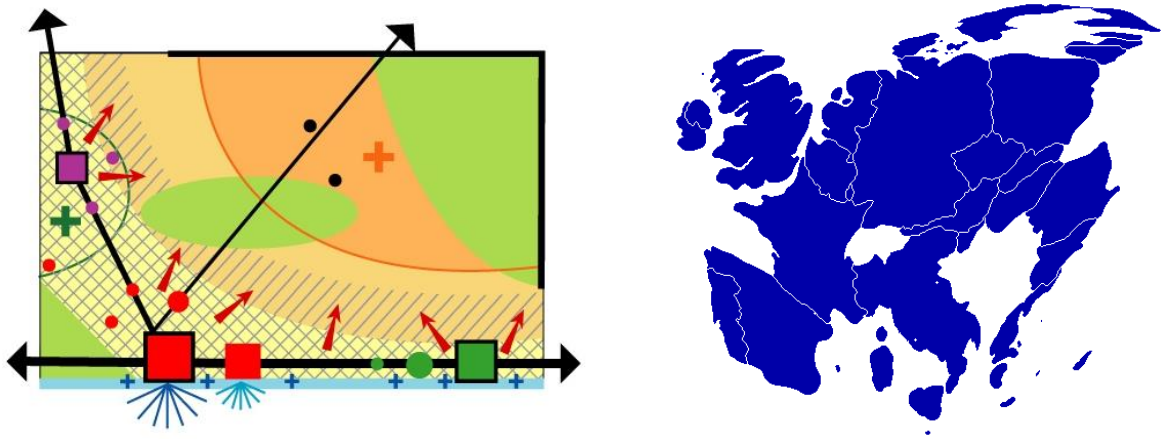
Cette « carte figurative», telle qu'elle est nommée par Minard, se trouve en fait être composée de plusieurs schémas et diagrammes, afin de représenter de très nombreuses variables dans le temps et l'espace : les nombres d'homme dans l'armée lors de l'aller et du retour (diagramme de Sankey), le contexte géographique des déplacements, le temps, la distance et la température. Cette représentation des effectifs de l'armée Napoléonienne est donc utilisée afin de présenter un phénomène spatio-temporel, qui met en avant les importantes pertes de l'armée. Mais avant tout, cette carte démontre un important changement dans le rôle de cartographier l'information : en plus

de communiquer des informations quantitatives et spatiales, Minard permet ici de mener une analyse de ces informations et nous propose de comprendre pourquoi l'armée a vécu tant de pertes. L'explication peut venir des conditions extrêmes de températures ou de l'environnement (pertes importantes aux passages de fleuves). Cette carte montre alors un grand pivot dans le rôle des cartes et autres méthodes de représenter l'information.

Poussée par l'essor des technologies de l'information (télécommunication, informatique, Web, Cloud, etc.) la cartographie a ensuite dû s'adapter à de nouveaux modes de diffusion de la donnée, de nouveaux supports et de nouveaux besoins centrés sur l'utilisateur / acteur. Aujourd'hui, une carte peut être interactive, animée, participative, représenter un phénomène en temps réel, etc. Ainsi, ce qui auparavant servait à communiquer une information précise et étudiée a évolué pour devenir l'instrument premier de recherche d'informations. Ce n'est donc plus seulement le mode de communication qui a changé, mais le rôle même de la carte (DiBiase et al. 1992 ; MacEachren 1995 ; MacEachren et Kraak 1997).

Dans les années 90, les manières de représenter l'information ont dû, elles aussi, évoluer selon le besoin et la technologie ; mais non pas en allant vers une modélisation plus « réaliste » des phénomènes, mais plus abstraite. Par abstraction, nous signifions la perte de la réalité topographique pour aller vers un espace plus orienté sémantique : par exemple, l'utilisation de cartogrammes (cartes déformées sur la surface ou la distance, selon une variable thématique et quantitative) ou bien de chorèmes (cartes schématiques). La Figure 2-8 présente ces types de cartes abstraites. Un exemple type de carte abstraite, pourtant utilisée par un grand public de manière simple et efficace, est la carte du métro. Son but d'utilisation n'est pas de mesurer la distance réelle entre deux points, mais seulement d'indiquer comment se rendre d'un point A à un point B. Dans un article sur le rôle de l'animation en cartographie, DiBiase et al. (DiBiase et al. 1992) citent justement Muehrcke sur le rôle fondamental de l'abstraction en cartographie :

> *It is abstraction, not realism, that gives maps their unique power. (Muehrcke 1990)*



(a) (b)

Figure 2-8. Cartes abstraites : (a) chorème de la région PACA (*source : Magali Brunet 2013*) et (b) cartogramme de surface de la population dans les états membres de l'UE (*source : Wikimedia*)

Les différents besoins qui justifient l'utilisation de cartes peuvent être classés en quatre groupes (Brunet 1987 ; Béguin et Pumain 1994 ; Kergomard 2009). Ces buts d'utilisation ont essentiellement été proposés avant l'arrivée des Systèmes d'Information Géographique (SIG), et représentent ainsi essentiellement les besoins relatifs aux cartes statiques et non interactives. Ces besoins sont :

- Observer, Situer, Placer, Trouver
- Classer, Comparer, Evaluer
- Combiner, Interpréter, Comprendre
- Révéler, Modéliser, Structurer

Néanmoins, l'utilisation d'outils interactifs permet à différents publics d'accéder à l'information géographique : les besoins en termes de cartes ne dépendent plus seulement du but d'utilisation, mais doivent considérer de nombreux autres paramètres. De manière à caractériser l'utilisation et le rôle de la carte, Alan M. MacEachren décrit cette utilisation au sein d'un espace à trois dimensions : il appelle ce principe le « cube cartographique » ou ***(Cartographie)^3***. Dans cet espace, la visualisation d'informations est décrite comme complémentaire à la communication d'informations. L'utilisation de la carte et de l'information géographique dépend, selon lui, des trois facteurs suivants :

- **L'interaction** permise par la carte : selon les outils utilisés et le support de la carte, l'utilisateur peut interagir et explorer les informations (SIG, outils d'analyse, etc.), ou seulement les lire sans possibilités de manipulation (plan papier).
- **L'usager** et son mode d'utilisation de l'information : certaines cartes sont produites de manière à être diffusées et lues par un maximum d'audience, telles que les cartes topographiques ou les portails d'information géographique (Géoportail de l'IGN, Google Maps, etc.). Mais elles peuvent aussi n'être utilisées que par une personne unique, privée, qui mène une exploration spécifique.
- La **connaissance de l'information et des motifs** : l'information peut être connue, maîtrisée et expliquée, auquel cas la carte est utilisée dans un but de communication. Autrement, l'information doit être explorée et analysée afin de découvrir de possibles motifs.

La Figure 2-9 présente les trois axes du cube de la (Cartographie)^3 proposé par MacEachren, ainsi que la place de la visualisation et de la communication dans cet espace. En 1997, MacEachren et Kraak rappellent alors l'importance de ces rôles complémentaires de la visualisation et de la communication, qui ne doivent pas être vus comme deux processus distincts. Ceci doit être pris en compte lors de la réalisation de nouvelles méthodes pour la représentation et la manipulation d'information géographique.

> *« All map use involves both visualization (defined loosely as the prompting of visual thinking and knowledge construction) and communication (defined loosely as the transfer of information), but map use can differ considerably in which of these activities is emphasized. » (MacEachren et Kraak 1997)*

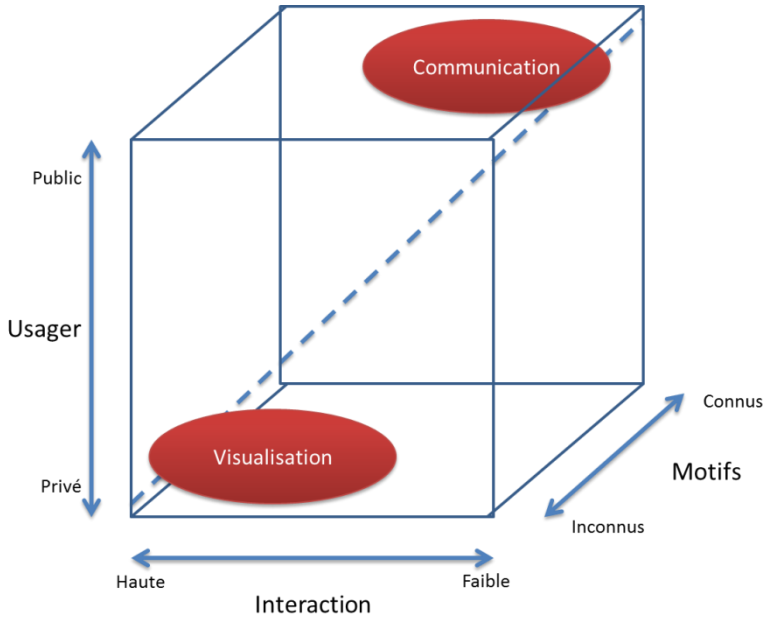


Figure 2-9. L'espace des utilisations de la carte : (Cartographie)[3] (d'après MacEachren, 1995)

La notion d'interactivité avec l'information et les outils qui la supportent a fortement fait évoluer la manière de représenter l'information visuelle et de l'utiliser. D'une carte qui était alors statique sur un format papier, nous sommes aujourd'hui arrivés dans une période où chacun peut créer sa propre carte selon ses intérêts et besoins, modifier les objets qui la composent, la manipuler, changer d'échelle, d'orientation, etc. Les normes qui géraient la cartographie statique et les règles pour « créer une bonne carte » ne sont aujourd'hui plus valables. On privilégie notamment le nombre croissant de données consultables et le nombre d'outils pour les analyser et les manipuler.

La prochaine sous-partie présente donc les grands courants de recherche menés dans la cartographie dite « interactive », et les nouvelles solutions pour visualiser et utiliser de tels outils.

2.1.2.2. L'évolution de la « carte interactive »

Grâce aux nouveaux outils informatiques utilisés pour la cartographie, cet art a de plus en plus évolué vers une vraie science informatique, de l'image, de l'interaction et de l'animation. Ainsi, les informations qui sont étudiées et représentées ne sont plus seulement utilisées à but de communication, mais aussi d'exploration, comme le montre le cube de MacEachren. Ces données sont alors spatiales, sémantiques, et temporelles (Peuquet 1994).

C'est ainsi que plusieurs recherches ont été menées sur la représentation du temps et du changement sur des cartes planes, grâce à des outils interactifs (Chen et al. 2008 ; Goldsberry et Battersby 2009 ; Andrienko et al. 2010). Le défi relevé par ce domaine de recherche s'intéresse à représenter de manière simple et efficace une information qui change dans le temps, grâce à des cartes choroplèthes animées par exemple (changement sémantique) (DiBiase et al. 1992 ; Goldsberry et Battersby 2009 ; Wilkening et Fabrikant 2011), ou encore par l'animation d'objets géométriques (changement de forme, position) (MacEachren et DiBiase 1991 ; Hurter et al. 2009 ; Hurter 2010). Ces recherches soulignent la grande difficulté de surpasser les limites cognitives de la vision et de la mémoire humaine, essentiellement à court terme.

En effet, comme nous l'avons présenté en début de ce chapitre, le système visuel humain impose de très nombreuses limites lors de la visualisation d'une séquence animée, comme le déplacement d'objets sur un réseau, ou l'évolution temporelle d'une carte choroplète. Ce problème concerne notre perception du changement : à cause de la multitude d'objets à étudier et des changements inopinés et inconnus de l'utilisateur, percevoir ces changements devient une tâche très complexe (Fabrikant et Lobben 2009 ; Goldsberry et Battersby 2009).

Ces limites cognitives sont d'autant plus accentuées lorsqu'un utilisateur observe un système animé, et qu'une mission spécifique lui est confiée : par exemple, repérer un certain type de données. Dans ce cas, les objets qui ne font pas partie du domaine d'intérêt (selon l'objectif) sont automatiquement effacés du champ de vision de l'utilisateur et oubliés. Ce problème a été mis en évidence grâce à la fameuse expérience du « Gorille Invisible » (Simons et Chabris 1999). La visualisation d'information changeante requiert donc d'étudier des méthodes spécifiques qui permettent de surveiller un type d'objet donné, sans perdre de vue l'état des objets voisins.

Parmi les manières possibles pour limiter cette erreur cognitive humaine, la notion de contrôle ou d'interaction est alors une nécessité à prendre en compte. La plus simple des solutions pour déjouer nos erreurs de compréhension d'une animation est, par exemple, de contrôler l'animation par des arrêts sur l'image et une reprise de l'animation (*play*, *pause*, etc.). L'ensemble des outils pour contrôler et interagir avec les informations affichées, représente alors une solution importante pour la compréhension de celles-ci.

L'animation et l'interaction sont alors deux notions distinctes, mais complémentaires, dans l'affichage et l'étude d'informations. Il est donc important de différencier la cartographie animée, ou dynamique, et la cartographie interactive : celles-ci se basent sur des données et outils différents. Les différents aspects de la cartographie « multimédia » sont résumés sur la Figure 2-10, ainsi que les différents niveaux de cartographie dynamique. Ce schéma, d'après les travaux d'Antoni et al. (Antoni et al. 2004), met en avant les différentes variables qui peuvent être ajoutées à une « carte d'affichage » simple, afin de créer ce que les auteurs appellent une « carte multimédia ».

Lorsque la dimension temporelle est prise en compte dans un affichage dynamique, plusieurs niveaux d'interactivité, ou de contrôle, peuvent être proposés. Une simple séquence animée qui passe en boucle afin de mettre en avant l'évolution d'un phénomène, ne propose qu'un très bas niveau d'interactivité. Ce cas de figure se présente plutôt dans le cas de cartes pour communiquer, représenté dans le coin inférieur du cube de MacEachren (peu d'interactivité). Une séquence animée peut aussi être rendue interactive par des contrôles et des informations supplémentaires ; ou bien générée en temps réel, telle qu'une carte du trafic maritime ou aérien. Ce dernier cas demande alors des interactions très avancées afin de pouvoir comprendre ces informations.

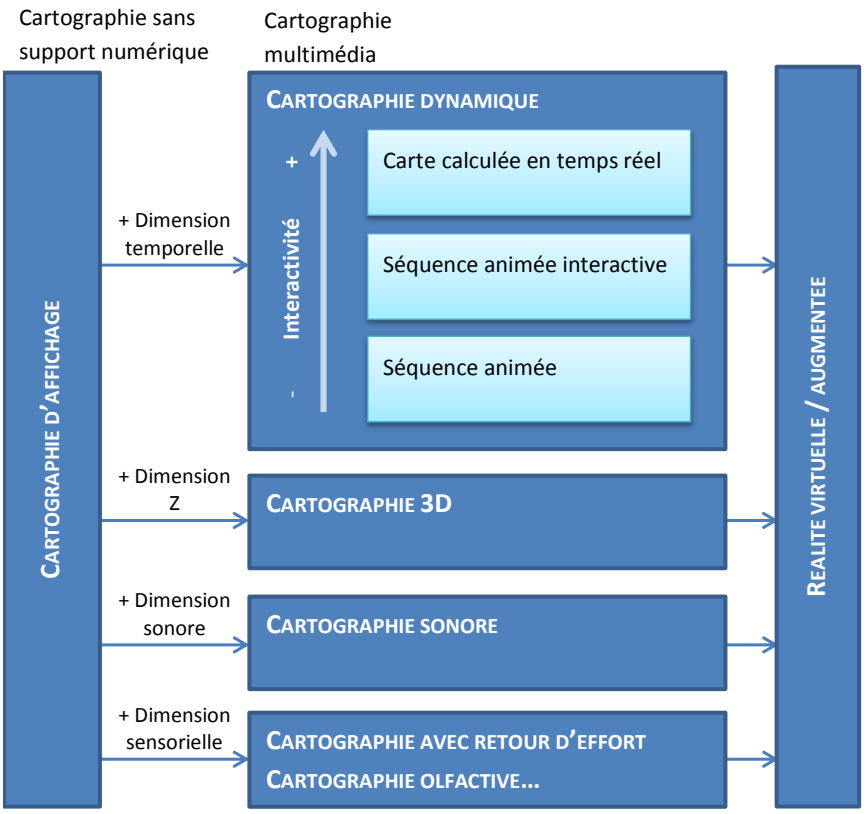


Figure 2-10. Cartographie multimédia (d'après Antoni et al. 2004)

Afin d'adapter la cartographie à ces nouveaux modes de représentation, de nombreux chercheurs ont proposé de nouvelles manières de représenter le changement. Parmi ces propositions, l'une conservait encore l'aspect statique des cartes, mais utilisait une représentation multiple permettant ainsi la comparaison de plusieurs états successifs d'une même carte (Monmonier 1990), ou bien une animation de ces états successifs dans un ordre défini (DiBiase et al. 1992). Cet ordre peut être le temps (représentation la plus simple à comprendre), ou bien être une variable attributaire ordinale (le temps se retrouve alors affiché sur la carte).

D'autres travaux ont porté sur la façon d'animer la carte, en complétant la liste des variables visuelles de Bertin (Bertin 1967) par de nouvelles variables d'animation et dynamiques. Parmi les travaux les plus importants en visualisation d'information dynamique, nous pouvons citer les recherches de Hayward (Hayward 1984) pour les variables de l'animation (voir la

Table 2-2), puis les recherches de Kraak et Ormeling (1996) pour les outils de la cartographie interactive (qui sont aujourd'hui les plus courants dans les SIG et portails cartographiques). Les travaux de MacEachren (MacEachren 1995) et DiBiase (MacEachren et DiBiase 1991 ; DiBiase et al. 1992) viennent ensuite compléter les résultats de Hayward en introduisant les variables dynamiques de la cartographie.

Table 2-2. Variables de l'animation et dynamiques

Variable	Auteurs de référence
Animation	
Taille	(Bertin 1967 ; Hayward 1984)
Position	(Bertin 1967 ; Hayward 1984)
Orientation	(Bertin 1967 ; Hayward 1984)
Vitesse de la scène	(Hayward 1984)
Couleur	(Bertin 1967 ; Hayward 1984)
Texture	(Bertin 1967 ; Hayward 1984)
Perspective (point de vue)	(Hayward 1984)
***Shot* (distance)**	(Hayward 1984)
Son	(Hayward 1984 ; Krygier 1994)
Dynamique	
Durée	(MacEachren et DiBiase 1991 ; DiBiase et al. 1992)
Ordre	(DiBiase et al. 1992)
Taux de changement	(DiBiase et al. 1992)
Synchronisation	(DiBiase et al. 1992)

Aujourd'hui, l'utilisation de nouvelles technologies à différentes échelles, telles que les smartphones, les tablettes tactiles, les murs de données, les cubes immersifs, etc., requièrent encore de nouvelles variables, de nouveaux outils et de nouvelles visualisations qui soient mieux adaptés aux dispositifs et aux besoins des utilisateurs (Tominski et al. 2011). Le changement des supports de la carte et de la manière d'interagir avec celle-ci a donc provoqué de grands changements dans notre rapport avec l'information géographique. Le rôle de la carte, et de la visualisation en général, ont pu évoluer vers l'analyse de l'information. Dans la sous-partie suivante, nous présentons donc le concept de visualisation d'information, son évolution vers l'analyse visuelle, ainsi que le concept central pour cette thèse : l'analyse géovisuelle.

2.1.3. La visualisation d'information : explorer pour découvrir l'inconnu

L'utilisation de supports variés et de données hétérogènes, à des fins divers de la communication à l'exploration de données, souligne alors l'intérêt d'ouvrir notre champ d'intérêt. Pour cela, il est nécessaire de voir beaucoup plus loin que la seule carte. Dans cette optique, le domaine de la visualisation d'information a alors été proposé au cours des années 2000, afin de permettre une interactivité beaucoup plus importante pour la manipulation d'information visuelle. Dans cette partie du chapitre, nous présentons ce domaine de recherche, aussi appelé InfoVis, et son application à l'information géographique, grâce à la géovisualisation.

2.1.3.1. Les domaines de l'InfoVis

La **visualisation d'information**, ou InfoVis (pour *Information Visualisation* en anglais), est un domaine très vaste de la recherche, issu de travaux communs entre l'infographie, l'interaction homme-machine (ergonomie) et la psychologie. La visualisation d'information s'intéresse notamment à la représentation d'**information abstraite**, c'est-à-dire lorsque l'information étudiée n'a pas de réalité

physique concrète, et sa représentation spatiale est choisie (Card et al. 1999). Une définition de la visualisation d'information est donnée par Card et al. en 1999.

> *The use of computer-supported, interactive, visual representations of abstract data to amplify cognition. (Card et al. 1999)*

Plus qu'un simple affichage d'information, la visualisation d'information s'inspire donc de nos capacités de perception (comme vu dans la sous-partie précédente) afin de mieux représenter et comprendre l'organisation de l'information dans des structures visuelles (Ware 2008 ; Ware 2000). La visualisation d'information est un domaine de recherche qui s'intéresse à des méthodes de représentation visuelle, qui nous permettent de **comprendre** et de **raisonner** à partir d'information visuelle, grâce à ses propriétés visuelles (Card et al. 1999 ; Heer et al. 2010). La Figure 2-11 présente le mode de fonctionnement de la cognition humaine, dans le cas du raisonnement et de la prise de décision : l'information perçue y est schématisée comme un stimulus externe, qui aboutit à une réponse (décision) à la fin du processus (Patterson et al. 2014). Sur cette figure, la cognition humaine est divisée en quatre composantes : la mémoire de travail, la reconnaissance de motifs, la mémoire à long terme et la prise de décision.

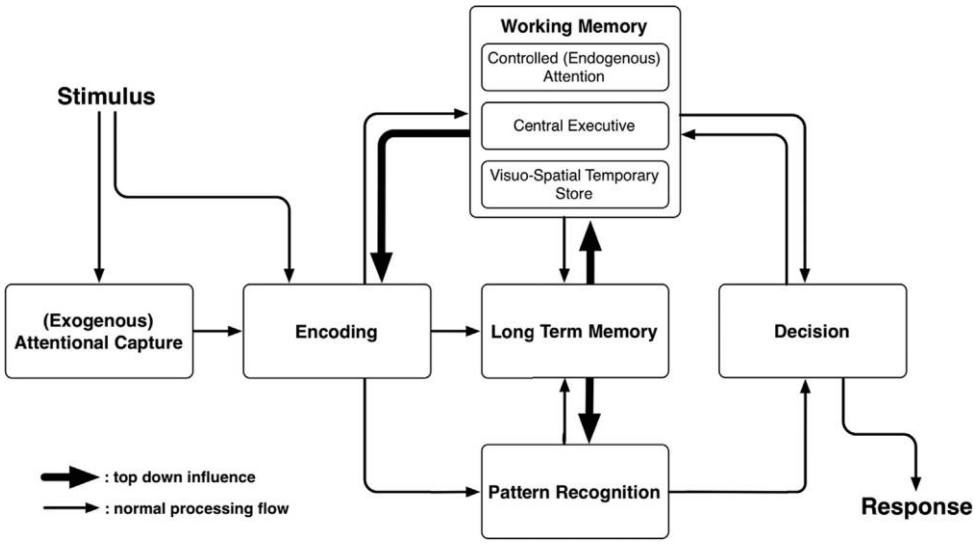


Figure 2-11. Schéma de la cognition humaine, dans le cas du raisonnement et de la prise de décision (*source :* Patterson et al. 2014)

Parmi les chercheurs qui ont participé aux fondations de ce domaine, nous pouvons retrouver les travaux de Tufte (Tufte 1983 ; Tufte 1997), de Bertin (Bertin 1967), de Card et Mackinlay (Mackinlay 1986 ; Card et Mackinlay 1997 ; Card et al. 1999). Ces travaux ont porté à la fois sur la théorie de la visualisation et des variables visuelles, comme présentées précédemment, et sur la proposition de nouvelles méthodes de visualisation.

Afin de décrire les étapes successives de la visualisation, la communauté scientifique emploie régulièrement le terme de *visualization pipeline* (Haber et McNabb 1990 ; Card et al. 1999 ; Chi 2000 ; Patterson et al. 2014). Ce processus est composé de quatre ou cinq étapes, qui traduisent la projection de l'information abstraite vers un espace visuel concret (Fabrikant et Skupin 2005). Ces étapes, présentées sur la Figure 2-12, sont les suivantes. Nous reviendrons plus en détail sur ces étapes dans la partie 2, traitant de la formalisation de la visualisation.

- Données brutes
- Tables de données, comportant des données filtrées ou non
- Projection visuelle, géométrique, de l'information
- Rendu final

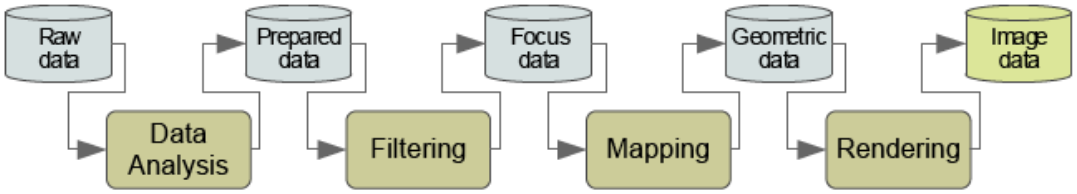


Figure 2-12. Processus de visualisation (*source : InfoVis-Wiki*)

Patterson et al. (Patterson et al. 2014) proposent ensuite une évolution de ce modèle du processus de la visualisation, en intégrant le processus de cognition humaine dans une boucle rétroactive (voir Figure 2-13). L'introduction de cette nouvelle étape sert alors à la fois de contrainte pour le design de la visualisation d'information (*defines*), ainsi que de retour d'utilisation afin de modifier ou changer la visualisation utilisée (*refines*). Selon les auteurs, avant de proposer une nouvelle visualisation, la cognition humaine doit donc être étudiée avec grand soin, de manière à comprendre le mode de fonctionnement de la compréhension et le raisonnement humain face à ces interfaces. Les principales finalités de la visualisation sont donc, d'après eux : faciliter la création de modèles mentaux de l'information, pour une meilleure utilisation de la mémoire à long terme, et une utilisation efficace des interactions pour réorganiser l'information selon l'importance des détails, pour un meilleur encodage visuel (Patterson et al. 2014).

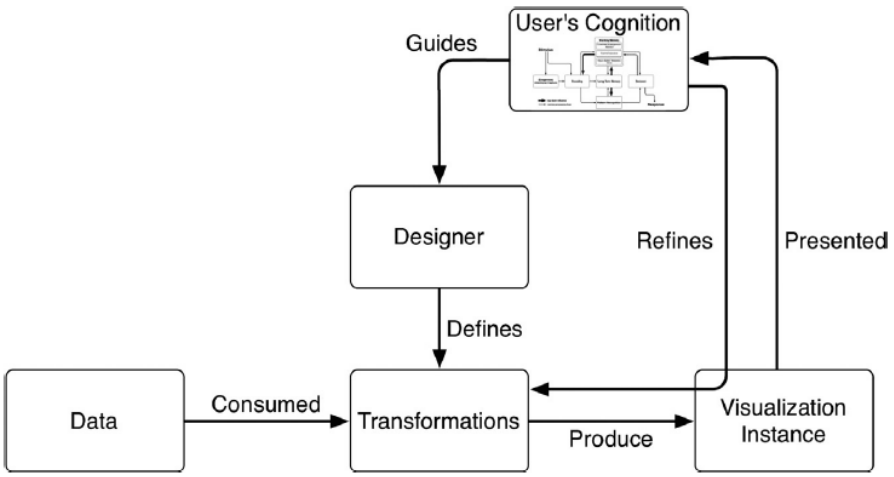


Figure 2-13. Processus de conception pour la visualisation d'information (*source :* Patterson et al. 2014)

Les applications de la visualisation d'information peuvent alors prendre la forme de visualisations simples telles que les graphiques et diagrammes, pour des variables à 2 ou 3 dimensions. Des formes plus avancées de la visualisation sont utilisées afin d'étudier les informations multidimensionnelles : les cartes d'arbres (*tree map*), les cartes de Kohonen (cartes auto-adaptatives, SOM) ou encore les coordonnées parallèles. Selon Thomas et Cook (2005), la visualisation d'information a pour but de *communiquer des informations abstraites de manière intuitive, grâce à notre capacité de voir, explorer et comprendre de grandes quantités d'informations*.

Le cas le plus fréquent de l'utilisation de l'InfoVis se retrouve dans l'étude de réseaux (informatiques, sociaux, etc.), qui peuvent se représenter sous formes de graphes, puis visualisés de très nombreuses manières. L'ouvrage de Lima (Lima 2013) présente une collection complète des très nombreuses méthodes pour la visualisation de réseaux complexes, et montre ainsi la grande diversité technique et artistique de ce domaine de recherche.

L'InfoVis se décline alors en plusieurs domaines spécialisés, tels que la visualisation scientifique et la visualisation d'information géographique. Dans le cas de la **visualisation scientifique** (SciVis), les informations utilisées et visualisées sont concrètes et leur représentation spatiale est alors contrainte par leurs propriétés physiques (Haber et McNabb 1990). Ce domaine de visualisation s'intéresse particulièrement à la représentation de phénomènes physiques pour la modélisation des matériaux et des forces qui s'y appliquent. Dans cette thèse, nous n'exploitons pas ce domaine, qui a sa propre communauté et des problématiques distinctes.

Enfin, la visualisation d'information géographique (GeoVis) se base, quant à elle, sur des données géoréférencées (MacEachren 1995). Elle vient de l'évolution des méthodes pour cartographier et analyser l'information géographique, que nous présentons plus en détail dans la sous-partie suivante.

2.1.3.2. La géovisualisation

La recherche en InfoVis, telle que présentée précédemment, est particulièrement utilisée pour la visualisation et l'étude de données abstraites (non concrètes). Les données et les connaissances portant sur l'espace géographique et le mouvement forment un tout autre ensemble de données, aux propriétés complètement différentes. Ainsi, leur visualisation a demandé l'instauration d'un nouveau concept qui suit l'InfoVis.

L'ensemble des techniques de visualisation des données à composantes géographiques forme donc le domaine de la **géovisualisation**. Ce domaine regroupe des méthodes et des outils qui peuvent aller du portail géographique le plus simple, aux systèmes de réalité virtuelle ou de réalité augmentée (Foyle et al. 2005) pour la représentation de données. Récemment, les interfaces SOLAP pour la fouille de données ont aussi été des propositions d'intérêt pour la recherche, en innovant dans le domaine de la modélisation de données spatio-temporelles et l'interaction pour la découverte et la compréhension de motifs. La géovisualisation consiste donc en la visualisation d'informations, telle que définie par Keim et al. (Keim et al. 2006), contenant des informations spatiales connues ou inconnues. Selon la description de ce domaine par Kraak (Kraak 2006), les cartes et les autres représentations visuelles de données doivent être manipulées dans un environnement interactif visuel, afin de stimuler la pensée visuelle et aider à la prise de décisions. Les travaux de Peuquet (Peuquet 1994), présentés dans le chapitre 1, soulignent l'intérêt de prendre en compte à la fois l'aspect spatial, l'aspect temporel et l'aspect sémantique des données afin de mieux comprendre les possibles motifs.

Kraak (Kraak 2006) regroupe les différents cas d'utilisation de la géovisualisation dans un cycle (Figure 2-14). Ce cycle peut alors être placé dans le cube de la (Cartographie)[3] de MacEachren, le long de sa diagonale. En effet, selon ces étapes, l'interaction avec les données, le type d'utilisateur et la connaissance des motifs vont évoluer. Le schéma de cette figure vient appuyer le fait que la carte ne doit pas être utilisée seule, mais en complément d'autres outils et visualisations. Au cours de ces

étapes, les cartes et autres visualisations viennent soutenir la pensée visuelle humaine, de manière à découvrir des motifs spatio-temporels dans la structure des données (Kraak 2006).

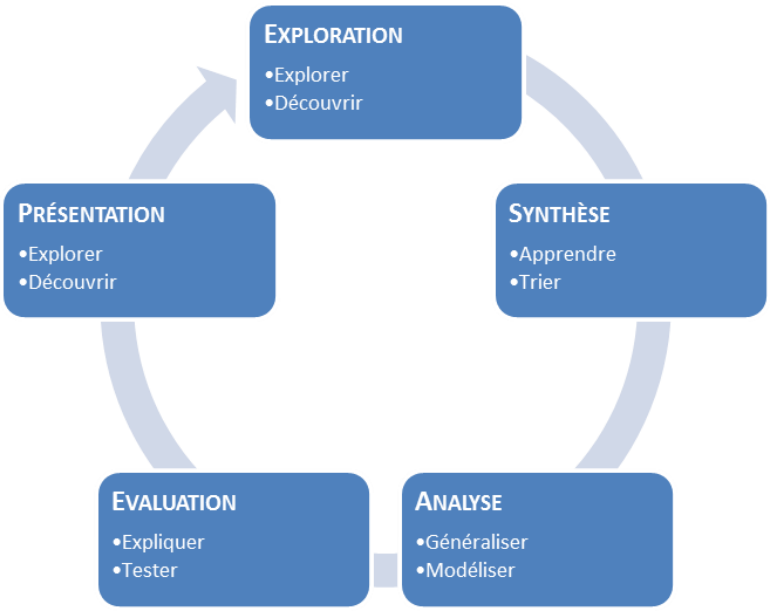


Figure 2-14. Rôles de la géovisualisation (Kraak 2006)

Aujourd'hui, le domaine de la géovisualisation a son propre groupe de recherche au sein de l'Association Cartographique Internationale (ACI), la *Commission on Geovisualization* présidée par Gennady Andrienko et Jason Dykes (ICA 2014). Cette communauté s'intéresse aux méthodologies d'analyse et d'exploration de l'information spatiale à partir d'interfaces visuelles interactives.

La pensée visuelle doit donc s'effectuer grâce à des représentations variées, mais en gardant un lien logique entre elles, stimulant ainsi la compréhension des phénomènes représentés. Afin de garder un lien entre ces différents modes de visualisation, des méthodes ont été mises au point, facilitant ainsi la recherche de correspondance sur l'ensemble des représentations. L'utilisation des variables visuelles et dynamiques communes (Bertin 1967 ; Hayward 1984 ; DiBiase et al. 1992 ; Kraak et Ormeling 1996) permet donc de relier plusieurs visualisations, grâce aux ressemblances ou différences des individus : une analyse multi-variée peut donc être effectuée sur plusieurs variables, plusieurs dimensions, etc.

2.1.3.3. L'analyse géovisuelle

Suite aux avancées du domaine général de la visualisation d'information, un nouveau terme est apparu dans la recherche : celui d'**analyse visuelle** (*visual analytics* en anglais). Ce terme a été introduit dans deux ouvrages fondamentaux : du côté de la recherche américaine, grâce aux travaux de Thomas et Cook (2005), et du côté de la recherche européenne avec les travaux de Keim et al. (Keim et al. 2006 ; Keim et al. 2010). Les deux définitions données pour ce nouveau concept sont les suivantes :

- la **science du raisonnement analytique**, facilitée par des interfaces visuelles interactives (Thomas et Cook 2005)

- la combinaison des techniques d'analyse automatisées et de visualisations interactives pour **supporter la compréhension, le raisonnement et la prise de décision** au sein de grandes quantités de données (Keim et al. 2010)

L'analyse visuelle est donc un domaine de recherche visant à proposer des environnements hautement interactifs, couplés à la puissance de calcul des machines, pour soutenir le raisonnement humain face à une grande quantité d'informations. Ces méthodes visuelles doivent ainsi supporter le pouvoir cognitif de l'utilisateur. Thomas et Cook proposent alors un mantra qui résume le rôle de l'analyse visuelle :

> *Detect the expected and discover the unexpected. (Thomas et Cook 2005)*

On retrouve ici deux notions principales : **l'analyse** d'information, qui permet de vérifier une hypothèse formulée (*expected*) ; et **l'exploration** de données et la recherche de connaissances au sein d'un ensemble de données non encore connues (*unexpected*). Par la suite, nous utilisons donc les termes d'analyse et d'exploration tels que nous les avons définis.

L'**analyse géovisuelle**, ou GeoVA (pour *Geovisual Analytics* en anglais) est alors la suite logique de l'utilisation de la géovisualisation à des fins de découverte de connaissances, d'exploration approfondie et d'analyse d'information à composantes géographiques. Lors d'une communication orale en 2012, Alan M. MacEachren a pu donner sa propre définition de ce domaine de recherche, qu'il formule ainsi :

> *Geovisual analytics focuses on visual interfaces to analytical methods that support reasoning with and about geo-information to enable insights about something for which place matters. (Kraak 2013)*

Une autre définition de l'analyse géovisuelle a pu être donnée récemment, suite au colloque GeoViz'13 à Hambourg, où Schiewe a pu regrouper les travaux présentés et synthétiser les aspects communs grâce à cette interprétation :

> *Geovisual analytics as a linkage of visual and computational methods and tools for extracting hypotheses and information from spatial data. (Schiewe 2013)*

Nous pouvons voir que les mots-clés *interface* et *interaction* sont fondamentaux dans ces définitions. Ils démontrent l'importance des interactions et de l'organisation des outils dans un environnement simple pour mener une analyse géovisuelle. De façon à pouvoir explorer des données dans toutes leurs dimensions, il est important de pouvoir visualiser et interagir avec chacune d'entre elles et de pouvoir basculer d'un espace à l'autre (Peuquet 1994 ; Kraak 2006). La Figure 2-15 donne un schéma d'environnement d'analyse et d'exploration pour l'analyse géovisuelle, tel que proposé par Kraak. Cet environnement basé sur les interactions entre cartes et autres visualisations de données doit être utilisé pour stimuler la pensée visuelle et aider à la prise de décision (Kraak 2006). Le but de l'analyse (géo)visuelle n'est donc pas d'apporter une réponse à une question, mais de **guider dans l'analyse et le processus intellectuel et personnel** que mène l'utilisateur (Roth 2012) grâce aux échanges d'informations entre l'interface et l'utilisateur (interactions).

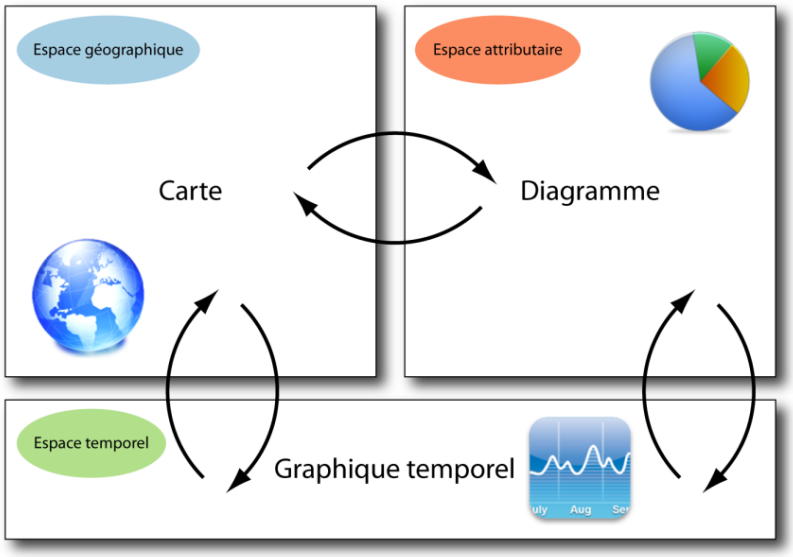


Figure 2-15. Environnement d'exploration pour l'analyse géovisuelle (d'après Kraak 2012)

2.1.3.4. Interagir entre les visualisations

Comme ont pu le mettre en avant les nombreuses définitions des concepts de visualisation et d'analyse visuelle, il est essentiel que ces méthodes proposent de nombreuses interactions avec l'information, c'est-à-dire un moyen de contrôler leur affichage et leurs propriétés. Shneiderman (Shneiderman 1996) pose un mantra connu sous le nom de **Visual Information Seeking Mantra**, qui résume les interactions clés en visualisation : ***Overview first, zoom and filter, then details-on-demand***. A ces interactions de base, il ajoute par la suite trois tâches d'analyse de l'information : *Relate*, *History* et *Extract*. Ces différentes actions doivent être permises au sein des visualisations, selon le type de données étudiées.

De manière à faciliter l'utilisation simultanée de visualisation au sein d'un même environnement d'exploration, une notion clé de l'interaction est le lien entre ces visualisations. En effet, en plus des nouvelles variables pour la représentation d'informations (variables dynamiques, variables de l'animation), de nouvelles variables et méthodes sont étudiées afin de faciliter le balayage visuel entre différentes visualisations et de relier rapidement les éléments communs entre ces visualisations. Un exemple simple d'interaction qui permet une manipulation rapide de la donnée est l'utilisation de filtres pour réduire la quantité de données affichées (Shneiderman 1983). L'utilisation de *sliders* est d'autant plus utile pour l'interaction avec des données temporelles linéaires (Kraak et al. 1997 ; Guo et al. 2011).

Mais l'utilisation de filtres demande de connaître au préalable la structure des données (attributs, valeurs, bornes, etc.). D'autres méthodes ont été proposées afin d'interagir avec la donnée de manière encore plus simple et sans connaître spécifiquement les valeurs recherchées.

La technique du *brushing* (élagage) (Roberts et Wright 2006) est une autre méthode, plus récente, d'interaction visuelle. Plusieurs plates-formes de visualisation, d'analyse visuelle et de géovisualisation proposent cette technique pour retrouver un élément ou un groupe d'éléments sur plusieurs visualisations à la fois. Cette technique permet d'isoler une entité ou un groupe d'entités du reste du jeu de données, sur une représentation (carte, tableau, graphe, etc.). Cette méthode peut aussi être utilisée de manière synchronisée sur plusieurs représentations : un lien visuel

(couleur, épaisseur, etc.) permet de relier simultanément les mêmes données sur plusieurs visualisations. La Figure 2-16 illustre la technique du brushing, appliqué à un diagramme en coordonnées parallèles. Sur cet exemple, chaque axe représente une variable quantitative (longueur, tirant, longitude, latitude, vitesse), et chaque ligne est un objet qui relie les valeurs de ses attributs. La sélection se fait par un cadre sur l'axe des longitudes : les lignes sélectionnées sont alors mises en avant, afin de rechercher un possible motif dans leur organisation.

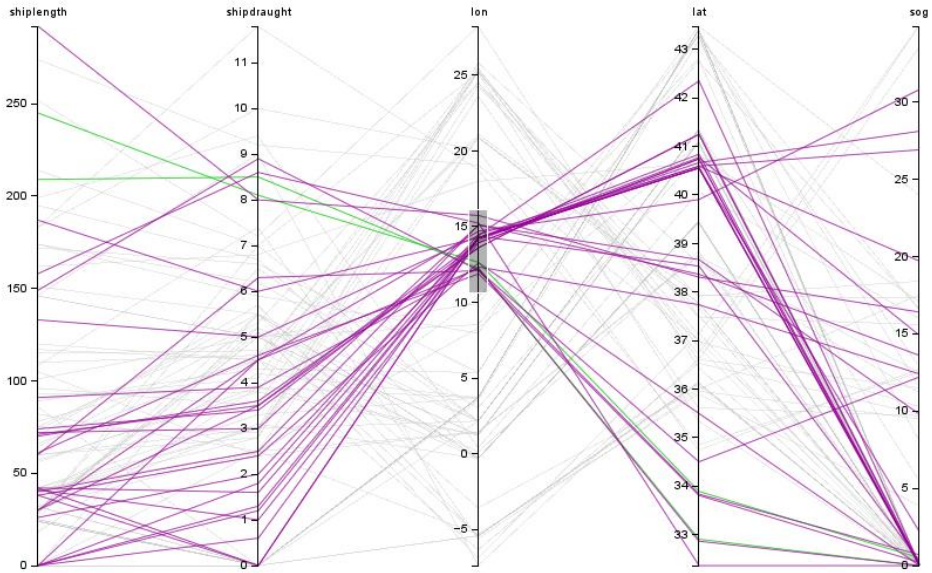


Figure 2-16. Technique du *brushing*, pour une sélection de longitudes moyennes (cadre de l'axe du milieu)

En 2011, Robinson propose la méthode du *highlighting* (surlignement, ou mise en valeur) qui affine le *brushing* (Robinson 2011) : il dégage d'abord les variables visuelles qui permettent de relier des éléments entre plusieurs visualisations (de type points, lignes, surfaces et texte), puis propose de nouvelles méthodes comme le contour, la *leader line* (qui relie le même élément sur plusieurs visualisations) ou la réduction de style qui permet d'avoir à la fois une vision locale (un élément) et globale (le voisinage, moins détaillé). La Figure 2-17 présente l'utilisation de plusieurs variables et techniques utilisées pour mettre en avant l'information sélectionnée, lors de l'utilisation simultanée de visualisations au sein d'un environnement d'analyse géovisuelle.

On retrouve ce mode de présentation dans les outils SOLAP (*Spatial Online Analytical Processing*) qui permet d'explorer de grandes quantités de données selon ces trois axes, tout en prenant en compte de nombreuses autres dimensions propres aux données étudiées (Rivest et al. 2001 ; Bédard et al. 2007). La Figure 2-18 donne un exemple d'interface utilisée dans un outil SOLAP, qui permet de synchroniser les nombreuses vues.

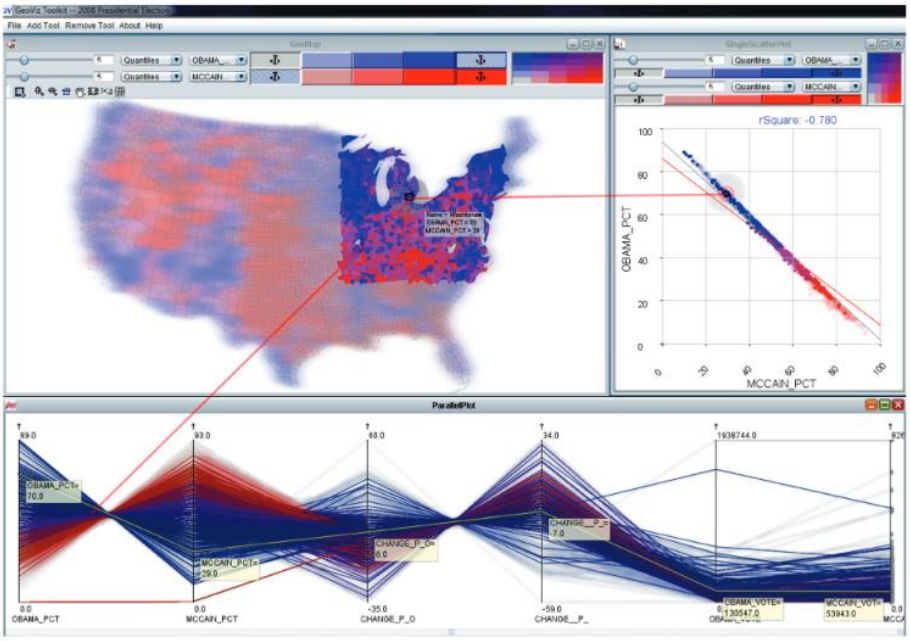


Figure 2-17. Utilisation de plusieurs liens entre les informations sélectionnées, sur plusieurs visualisations : flou, transparence, leader line (*source :* Robinson 2011)

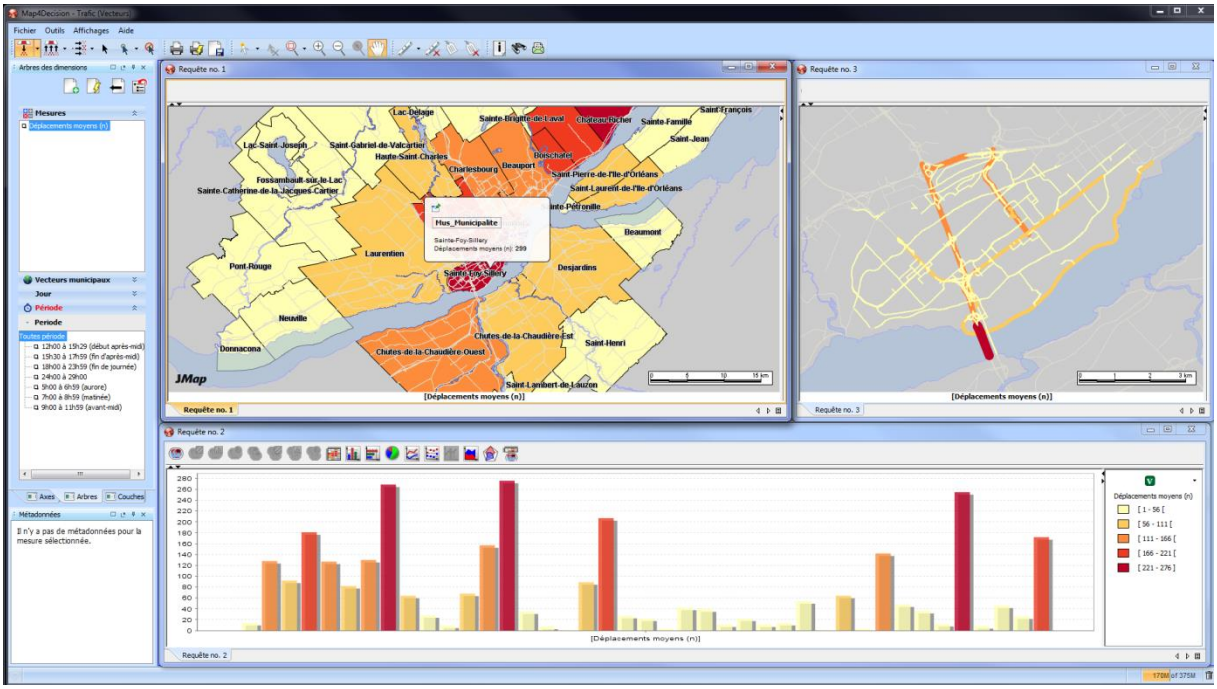


Figure 2-18. Interface de l'outil Map4Decision pour l'analyse SOLAP (*source : Intelli3*)

La communauté de l'analyse géovisuelle, et de l'analyse visuelle en général, étant très active, de très nombreuses méthodes sont développées, testées et évaluées chaque année pour la représentation et l'analyse de données spécifiques. Parallèlement à ces travaux, de nombreuses recherches ont pu être menées sur la modélisation de la visualisation, de manière à décrire et comparer les possibilités de ces nombreuses méthodes, de manière formelle. Dans cette première partie, nous avons expliqué

la théorie de la perception et de la visualisation, avec notamment l'ensemble des variables visuelles de Bertin. Dans la partie suivante, nous nous intéressons aux travaux qui, appuyés sur cette base théorique de la perception, ont pu introduire des modèles formels de la visualisation d'information, abstraite ou géographique.

2.2. Modélisation formelle des méthodes de visualisation d'information

Comme nous l'avons vu, Bertin a été parmi les premiers à proposer une théorie et une formalisation de la manière de représenter l'information visuelle (Bertin 1967). Depuis, de très nombreux travaux de recherche se sont intéressés à la modélisation des différentes étapes de la visualisation, selon les différentes étapes du pipeline présenté dans la partie précédente. Dans cette partie, nous présentons les travaux menés dans la formalisation des types de données à visualiser, des manières de les visualiser, et leur utilisation au sein de systèmes intelligents pour la visualisation automatique d'information.

2.2.1. Modélisation des données

2.2.1.1. Typologie de la donnée

Les premiers travaux relatifs à la modélisation de données remontent aux recherches de Stevens (Stevens 1946) et sa proposition d'échelle de données. Stevens met au point une échelle de quatre types de données, organisées selon leurs propriétés intrinsèques :

- **Données nominales :** l'identité de chaque donnée est propre, ce qui permet de la distinguer au sein d'un groupe.
- **Données ordinales :** les données peuvent être ordonnées et comparées. Cela induit que des données ordinales soient aussi nominales.
- **Données intervalles :** se basent sur des données quantitatives et réutilisent les propriétés des données nominales et ordinales. Des intervalles peuvent être égaux, différents, et introduisent la notion de recouvrement. Il n'y a pas de zéro absolu, celui-ci ne provient que de conventions.
- **Données ratios :** quatre relations sont possibles, l'égalité, l'ordonnancement, l'égalité d'intervalle et l'égalité de ratios. Les ratios introduisent la notion de zéro absolu. Des opérateurs mathématiques peuvent alors être utilisés pour la comparaison des données.

Ces types de données se retrouvent dans les niveaux des variables visuelles proposées par Bertin (Bertin 1967) dans ses travaux sur la sémiologie graphique (voir 2.1.1.3, p. 67). Les niveaux des variables visuelles dépendent du type de données que l'on affiche, mais surtout des opérations que l'on veut ensuite effectuer grâce à la visualisation. Des données ordonnées peuvent donc être seulement représentées par une variable visuelle de niveau nominal (ou qualitatif), si l'on juge que la visualisation ne doit pas servir à ordonner les informations.

Ware étant spécialisé dans le développement de visualisations, sa recherche met en avant le besoin de penser la structure des données en connaissant les problèmes de la visualisation. Il propose donc un modèle simple de données, mais beaucoup plus orienté visualisation et projection de l'information. Selon lui, les deux caractéristiques de base d'un ensemble de données sont leurs

valeurs et leur **structure** (Ware 2000). La structure des données représente la relation qu'elles ont entre elles, ainsi que ce qui est donc utilisé en visualisation pour mieux percevoir et comprendre les motifs émergents. Néanmoins, la valeur des informations doit elle-même être caractérisée, ce pour quoi Ware propose trois types identifiés par Stevens : **catégorie**, **entier** et **nombre réel**.

La Table 2-3 reprend ces trois études et compare les différents noms des groupes identifiés par Stevens, Bertin et Ware.

Table 2-3. Typologie des données

Stevens (1946)	Bertin (1967)	Ware (2000)
Nominal	Qualitatif	Catégorie
Ordinal	Ordonné	Entier
Intervalle	Quantitatif	Nombre réel
Ratio		

2.2.1.2. Typologie de l'information

Tandis que les types de données proposés par ces chercheurs s'intéressent particulièrement aux comparaisons des données entre elles, plusieurs études se sont intéressées à la typologie de l'information contenue dans les données et sa structure. C'est ce que Ware explique lorsqu'il propose d'anticiper la visualisation en étudiant les relations entre les informations utilisées. Ainsi, Shneiderman (Shneiderman 1996) propose une classification des informations selon les groupes suivants, qui influent fortement sur la manière de les représenter :

- **Unidimensionnelles :** données linéaires, texte, etc.
- **Bidimensionnelles :** données planaires telles que les points, lignes, polygones, ou encore les images.
- **Tridimensionnelles :** objets du « monde réel » et objets complexes 3D.
- **Multidimensionnelles :** données complexes, bases de données, etc.
- **Temporelles :** « lignes de vie » de variables qui évoluent dans le temps, définies par un début, une fin et de possibles recouvrement entre objets.
- **Arbres :** organisation hiérarchique, par des liens parents-enfants.
- **Réseaux :** organisation plus complexe qu'un arbre, avec des liens multiples.

Heer et al. (Heer et al. 2010) proposent une classification plus récente de la structure de l'information, selon les méthodes de visualisation avancées qui ont pu être proposées depuis les travaux de Shneiderman. Les classes identifiées sont alors :

- **Distribution statistique et Série discrète**
- **Série temporelle et Série continue**
- **Hiérarchie**
- **Réseau**
- **Carte**

A ces types d'informations s'ajoutent ensuite les très nombreux types de cartes, dont quelques-unes ont pu être présentées dans la partie 2.1.2, ainsi que les représentations symboliques (Kosslyn 1989),

très utilisées aujourd'hui à des fins de communication simple et efficace en infographie (voir Figure 2-19). Ces exemples d'infographie repoussent les limites du possible à une infinité et se retrouvent contextualisés au sujet présenté. Toutefois, ces modes de représentation de l'information ne peuvent pas être considérés comme de la « visualisation d'information » à proprement parler, car ce ne sont que des représentations statiques (images) et non interactives. Nous sommes donc dans le coin inférieur du cube de MacEachren, qui représente la communication d'informations connues auprès d'un large public.

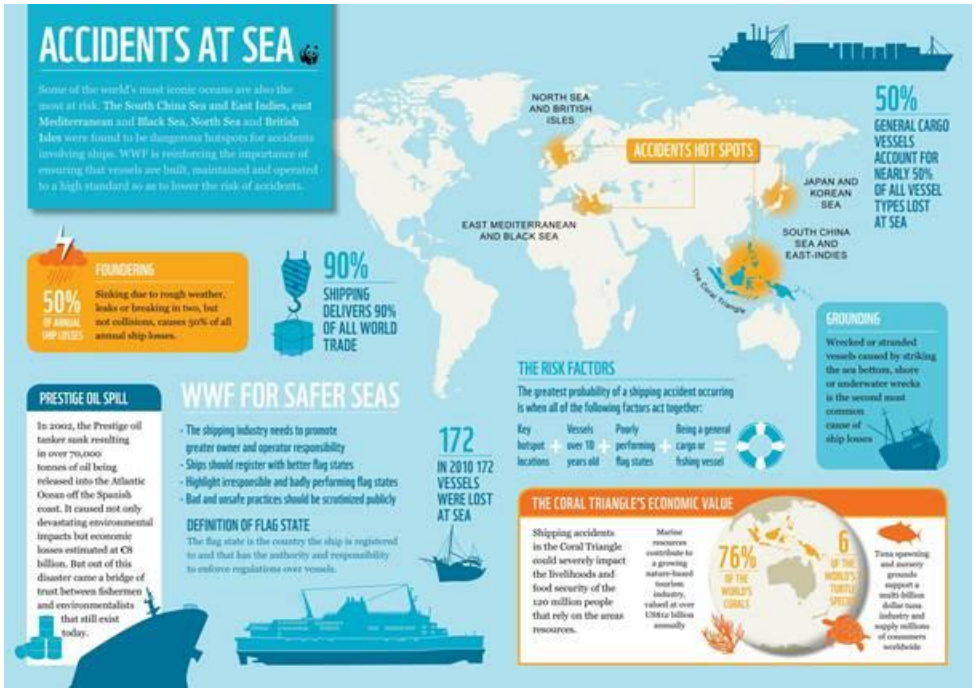


Figure 2-19. Infographie pour présenter les différentes zones d'accidents maritimes et les facteurs de risques (*source : www.maritimepassivesafety.com*)

De manière à reprendre ces typologies de l'information, la Table 2-4 résumé les différentes dénominations proposée par Shneiderman (1996), et Heer et al. (2010) pour cette classification. Keim (Keim 2002) ajoute le type *Algorithme* dans les informations à visualiser, mais cet aspect relève d'un tout autre domaine, purement InfoVis. Dans notre intérêt de l'analyse géovisuelle de mouvement, les algorithmes ne rentrent pas dans le type d'informations à visualiser.

Table 2-4. Typologie de l'information

Shneiderman (1996)	Heer et al. (2010)	
Unidimensionnel	Distribution statistique Série discrète	
Bidimensionnel		Cartes
Tridimensionnel		
Multidimensionnel		
Temporel	Série temporelle, série continue	
Arbre	Hiérarchie	
Réseau	Réseau	

Suite à ces très nombreux travaux sur la recherche d'une taxonomie des données et de la structure l'information étudiée, plusieurs recherches se sont intéressées à la manière d'utiliser et d'étudier ces information au sein de règles. Elles tendent ainsi à proposer des méthodes de visualisation automatisées selon le type de données étudiées et selon les règles de la représentation graphique. La prochaine sous-partie présente les différentes méthodes qui ont été proposées, se basant sur ces travaux de taxonomie.

2.2.2. Modélisation des méthodes de visualisation

De très nombreux modèles et classifications des méthodes de visualisation d'information ont été proposés, de manière à décrire le processus de transformation / projection de l'information vers un espace visuel. Ces modèles prennent en compte la dimension visuelle et les variables visuelles utilisées, mais parfois aussi les interactions et les transformations des données vers l'espace visuel (*visualization pipeline*).

2.2.2.1. Classification de Keim

Une classification simple des techniques de visualisation a été effectuée par Keim (Keim 2002), dans laquelle il identifie trois axes qui caractérisent une visualisation : les types de données visualisées, les interactions possibles et les techniques de représentation utilisées. Keim présente alors son modèle comme un espace à trois axes, tel que représenté sur la Figure 2-20. Pour cela, il utilise la typologie de Shneiderman, agrémentée du niveau supplémentaire « algorithme ». Les interactions proposées sur son axe dédié représentent les contrôles directs de manipulation de données, telles que la projection, le filtrage, le zoom, la distorsion et le lien (Kreuseler et al. 2000 ; Robinson 2011). Pour la classification des techniques de représentation de l'information, Keim utilise les classes suivantes :

- **Affichage standard 2D/3D :** graphiques, diagrammes, etc.
- **Affichage dont la géométrie est transformée :** coordonnées parallèles, etc.
- **Affichage d'icônes**
- **Affichage par densité de pixels :** matrices de données colorées, heatmaps, etc.
- **Affichage en piles :** cartes d'arbres, etc. Données hiérarchiques.

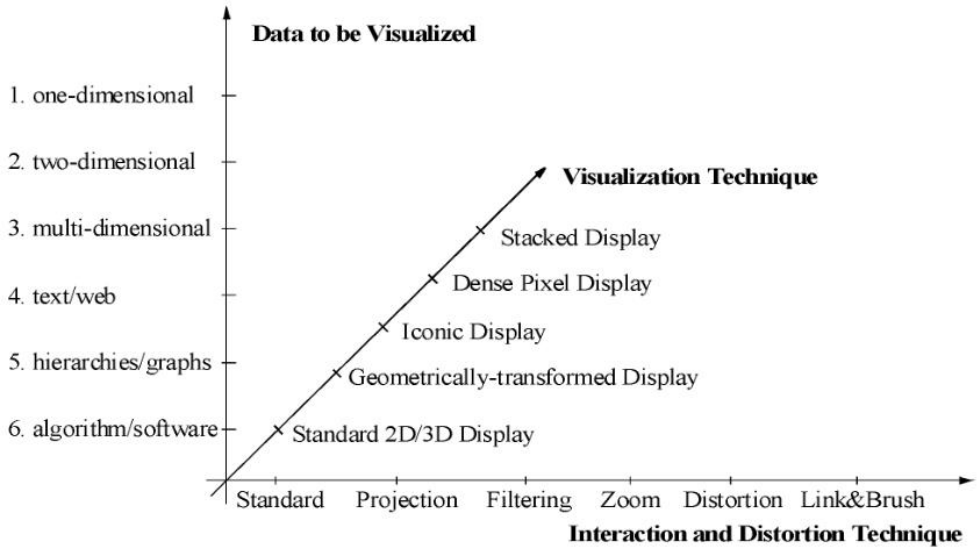


Figure 2-20. Classification des techniques de visualisation d'information (*source :* Keim 2002)

Toutefois, cette proposition se limite à une comparaison des grands types de visualisation et s'applique très mal au domaine de la géovisualisation. Cette recherche permet essentiellement de classifier les techniques de visualisation et d'interaction avec des données abstraites en InfoVis.

2.2.2.2. Modélisation formelle : le Reference Model Pattern

De manière à pouvoir modéliser des espaces de visualisation, il est nécessaire de prendre en compte différentes classes, qui permettent de caractériser les possibilités de la visualisation et le processus de transformation de l'information. La modélisation des méthodes de visualisation doit donc prendre en compte toutes les étapes de ce processus, de la donnée brute à la donnée visuelle. Pour cela, il est recommandé de faire une distinction claire entre les modélisations des données et les modélisations de l'espace visuel (Card et Mackinlay 1997). Cette distinction permet en effet d'améliorer, de corriger et d'adapter ces différents modèles de manière indépendante, selon leur domaine d'application et selon les besoins identifiés. Ce concept est connu comme le ***Reference Model Pattern*** (RMP) (Heer et Agrawala 2006), illustré sur la Figure 2-21. Ce modèle met en avant cinq classes nécessaires à la modélisation de la visualisation : la source de données brutes (*DataSource*), qui crée les tables de données (*DataSet*), l'espace de visualisation (*View*) et ses contrôles (ou interactions) et finalement la visualisation, qui est une représentation (projection) spécifique de tables de données choisies.

Depuis les années 90, de nombreuses recherches se sont basées sur ce modèle de référence pour proposer différentes modélisations de la visualisation d'information. Décrire les différentes étapes successives des données permet en effet de mieux décrire les **interactions possibles** avec l'information, ceci à chaque étape de la visualisation (Haber et McNabb 1990 ; Chi 2000). Cette modélisation générique est donc très avantageuse pour une description formelle des espaces de visualisation et de leurs possibilités.

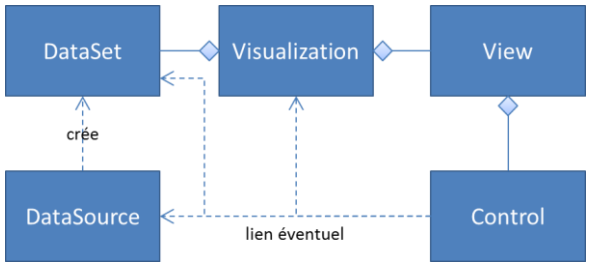


Figure 2-21. Reference Model Pattern pour la visualisation d'information (d'après Heer et Agrawala 2006)

Dans la suite de ce manuscrit, nous présentons deux modèles qui mettent en application le RMP, grâce aux travaux de Card et Mackinlay (Card et Mackinlay 1997), puis de Chi (Chi 2000). Ces travaux sont alors comparés et étudiés pour une utilisation dans un système intelligent de formalisation de connaissances.

2.2.2.3. Modèle de Card et Mackinlay

Les travaux de Card et Mackinlay (Card et Mackinlay 1997) ont abouti à une caractérisation des visualisations par une représentation tabulaire. Chaque ligne correspond à une information encodée visuellement, tandis que les colonnes représentent le processus de la projection vers un espace de visualisation : filtrage, transformation, projection dans l'espace visuel, les variables visuelles utilisées, puis les interactions possibles avec cette donnée. Ce modèle a ensuite été étendu par les travaux de Hurter et Conversy (Hurter et Conversy 2007) afin de prendre en compte l'affichage dynamique dans

ce modèle, tel que l'animation des objets (application à la visualisation du trafic aérien). Les travaux originaux de Card et Mackinlay ont eu pour but de comparer et étudier l'utilisation de l'espace pour le design de la visualisation, mais n'est pas adapté à une utilisation informatisée telle qu'une base de données. La Table 2-5 présente ce modèle, avec l'extension temporelle. Les colonnes proposées par Card et Mackinlay sont identifiées par une lettre, dont les significations sont les suivantes :

- D : type de données, selon la classification de Bertin (noté N, O, Q) ou plus complexe (Qx quantitative spatiale, QXlon géographique, NxN nominale bidimensionnelle)
- F : fonction ou filtre pour créer un sous-jeu de données
- D' : type de données après opération de la fonction, noté de la même façon que D
- X, Y, Z : position dans l'espace de visualisation
- T : position temporelle
- R : variable visuelle utilisée, propriété rétinienne
- - : connexion, lien entre les objets graphiques
- [] : lien d'encapsulation entre les objets graphiques
- CP : en cas de perception contrôlée, à l'inverse de la perception automatique (texte, chiffres)

Les colonnes Fr et A proviennent des travaux de Hurter et Conversy (Hurter et Conversy 2007) et correspondent au numéro de *frame* affichée à l'écran, et A pour les animations indépendantes des données, c'est-à-dire un changement de variable visuelle.

Table 2-5. Modèle de Card et Mackinlay, agrémenté de l'extension dynamique

Donnés				Perception automatique									
Nom	D	F	D'	X	Y	Z	Fr	A	T	R	-	[]	CP

Afin de compléter les cellules de la table, pour ce mode de représentation, un ensemble de notations est indiqué par les auteurs, de manière à unifier l'utilisation d'un tel modèle. Mais cette représentation n'a pas été utilisée dans un « langage machine » pour modéliser la visualisation de manière plus formelle et interopérable.

2.2.2.4. Modèle de Chi

Un autre modèle de grand intérêt a été proposé par Chi (Chi 2000), grâce au Data State Reference Model (DRSM). Tout comme le modèle de Card et Mackinlay, ce modèle suit le cadre du Reference Model Pattern en décrivant chaque étape de la projection des données vers un espace visuel. Ces étapes successives sont appelées des **abstractions**, séparées par des **transformations**, et permettent chacune des **opérations** à l'intérieur des étapes. Ces étapes (illustrées sur la Figure 2-22) sont, tout comme dans le RMP :

- **Valeur :** donnée brute
- **Abstraction analytique :** table d'information, méta-données
- **Abstraction visuelle :** information visualisable par l'œil
- **Vue :** finalité de la visualisation dans un espace manipulable

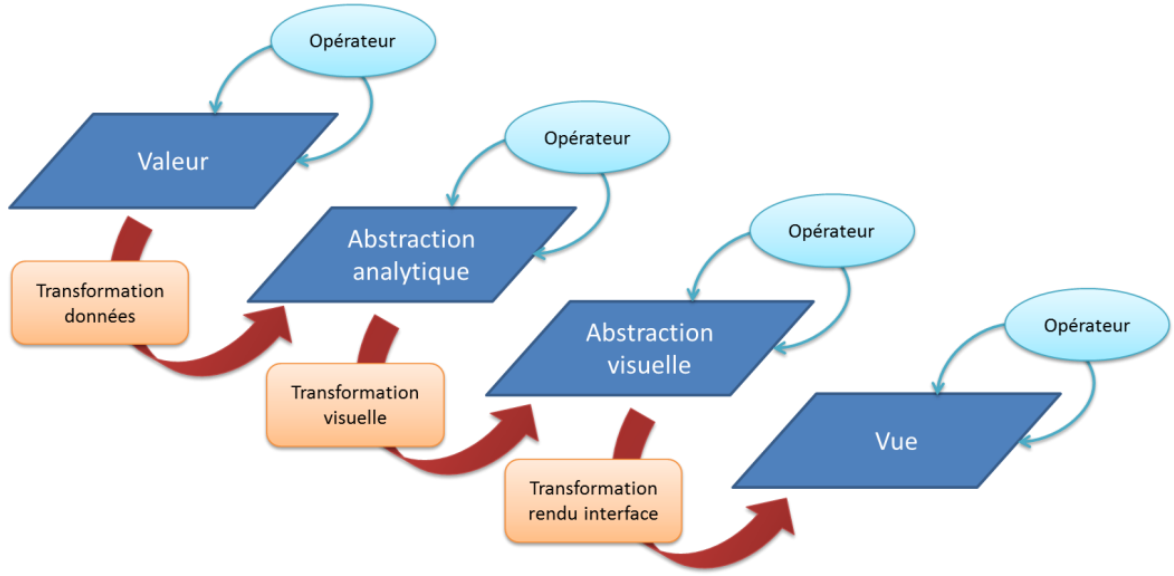


Figure 2-22. Data State Reference Model pour la modélisation de la visualisation d'information (d'après Chi 2000)

Chacune de ces étapes permet donc à l'utilisateur de contrôler la donnée, la modifier et contrôler la transformation vers l'étape suivante. Le DSRM était alors appliqué à la visualisation d'information (abstraite), domaine où la donnée ne contraint pas sa projection visuelle – contrairement à la visualisation d'information géographique.

Les travaux de Daassi et al. (Daassi et al. 2005) ont ensuite permis d'approfondir l'utilisation du *Data State Reference Model* pour une application à la visualisation d'information **temporelle**, c'est-à-dire qui évolue dans le temps, et qui requiert éventuellement une animation. De manière à décrire l'abstraction analytique puis l'abstraction visuelle de l'information, les auteurs proposent d'intégrer la dimension temporelle en plus de la dimension sémantique des données, dans ce processus de visualisation. Différentes stratégies sont testées afin d'en évaluer leur utilité. Le temps et la donnée sémantiques peuvent en effet être visualisés séparément (dans l'espace de vue), ou bien être intégrés lors d'une étape en amont de la visualisation : lors de l'abstraction analytique, ou bien visuelle. La Figure 2-23, issue de l'article par Daassi et al., illustre un exemple d'intégration des données au niveau de l'étape d'abstraction visuelle.

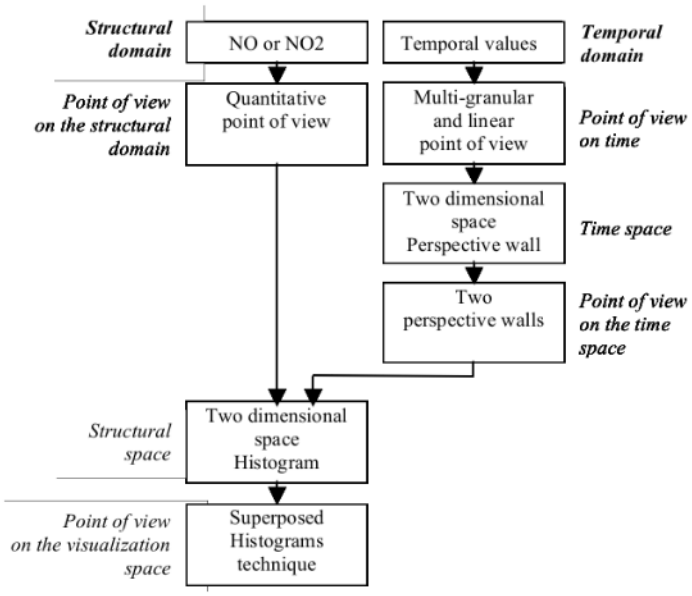


Figure 2-23. Modélisation du processus de visualisation pour des histogrammes superposés (*source :* Daassi et al. 2005)

Les auteurs étudient ainsi 37 outils de visualisation, basés sur la taxonomie qu'ils ont pu identifier. Certaines étapes d'intégration des données sémantiques et temporelles sont alors jugées inefficaces ou inutiles, telles que visionner le temps et la sémantique de manière complètement distinctes, ou bien fusionner les abstractions visuelles à la toute fin du processus (problème de compatibilité). Selon le modèle de Chi, ces travaux se basent sur les quatre étapes successices de la données brute (*Data*), l'abstraction analytique (point de vue donnée *PVD*), l'abstraction visuelle (*Space*), et le point de vue « espace de visualisation » (*PVS*). Les méthodes jugées comme les plus intéressantes sont l'utilisation de données intrinsèquement temporelle, la fusion dès l'abstraction analytique, ou bien la fusion lors de l'abstraction visuelle (voir Figure 2-24).

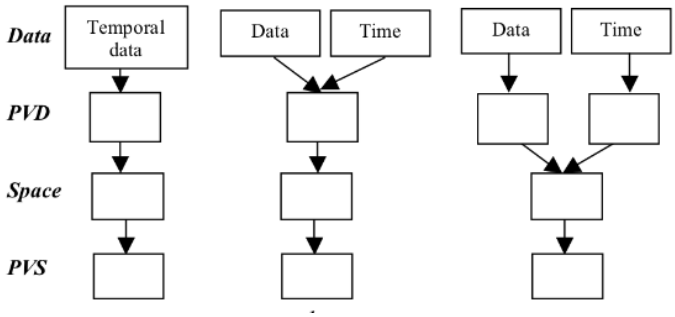


Figure 2-24. Intégration de la dimension temporelle en visualisation d'information (*source :* Daassi et al. 2005)

Cette modélisation des méthodes de visualisation de l'information a été testé sur plusieurs méthodes, dont certaines spécifiques à la géovisualisation, à partir de l'étude de méthodes fondamentales. Nous présentons certains de ces exemples dans les annexes de cette thèse, sur l'annexe *3. Modèle Data State Reference Model de Chi : exemples*, p.237.

2.2.2.5. Autres modèles « non RMP »

Plusieurs modélisations de la visualisation ont aussi été proposées, sans suivre le formalisme du Reference Model Pattern. Parmi ceux-ci, les travaux de Bertin (Bertin 1967) peuvent être cités

comme les premiers à proposer un modèle formel des possibilités de visualisation. Sa proposition a pris la forme d'un code par icones, selon les niveaux de l'information et les dimensions dans l'espace visuel (appelés *mobilisation des dimensions*). La Figure 2-25 présente les symboles utilisés par Bertin pour sa caractérisation de la visualisation.

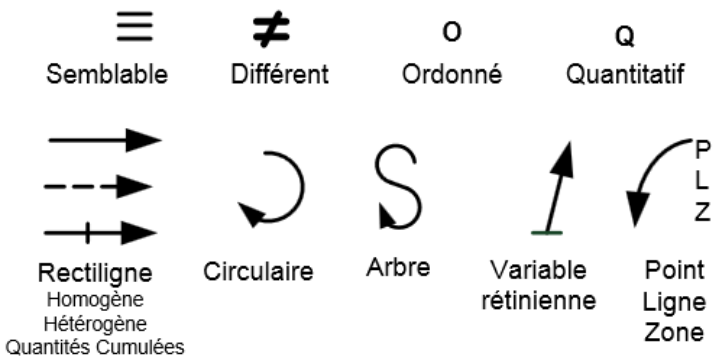


Figure 2-25. Symboles de Bertin pour la caractérisation de visualisations

Les visualisations telles que les graphes, les courbes, les histogrammes, etc., sont alors représentées par un assemblage schématique de ces différents symboles. Dans sa thèse, Hurter (Hurter 2010) propose plusieurs exemples d'utilisation de cette symbologie, dont la Figure 2-26 est extraite : il s'agit ici de l'utilisation d'une variable visuelle, rétienne, pour représenter un histogramme à axes parallèles. Toutefois, cette caractérisation est peu complète et n'est pas traduisible en langage machine pour une utilisation dans un système intelligent. Elle reste pourtant simple d'utilisation, notamment pour comparer plusieurs méthodes de visualisation par un être humain.

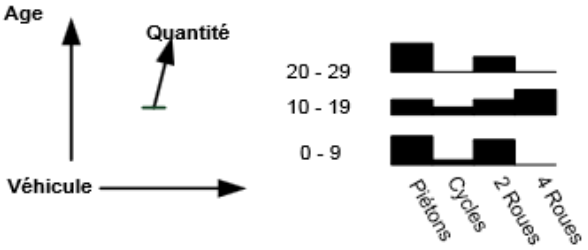
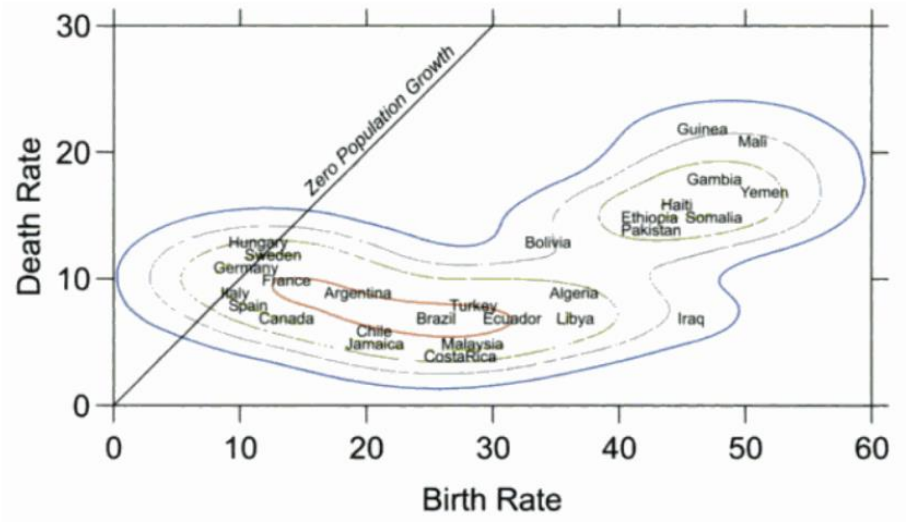


Figure 2-26. Schématisation de Bertin pour la caractérisation d'un histogramme à axes parallèles (repris dans Hurter 2010)

D'autres travaux de recherche ont optés pour une représentation textuelle de la visualisation, et non pas symbolique comme celle de Bertin. On peut citer les travaux de Wilkinson, ou encore ceux de Baudel, qui ne suivent pas non plus le principe du Reference Model Pattern.

Wilkinson (Wilkinson 1999) s'appuie sur sa propre interprétation du processus de visualisation, qu'il sépare en trois étapes successives et modélisables : la **spécification graphique**, l'**assemblage** et l'**affichage**. La spécification permet de décrire la transformation des données, l'échelle et le système de projection, les variables visuelles et enfin les aides visuelles pour contrôler la perception. L'assemblage décrit la structure et la disposition des objets dans l'espace de visualisation (scène), et l'affichage renseigne sur le support de la visualisation. La Figure 2-27 présente un exemple de caractérisation d'une visualisation, en deux dimensions, qui regroupe plusieurs projections : une première représentation scalaire par un nuage de points, et une seconde par noyau de Kernel.




```
ELEMENT : points(position(birth*death), size(0), label(country))
ELEMENT :
contour(position(smooth.density.kernel.epanechnikov.joint(birth*death)),color.hue())
GUIDE : form.line(position((0,0),(30,30)),label(« Zero Population Growth »))
GUIDE : axis(dim(1),label(« Birth Rate »))
GUIDE : axis(dim(2),label(« Death Rate »))
```

Figure 2-27. Caractérisation de Wilkinson pour l'étude du taux de mortalité en fonction du taux de natalité, ainsi que leur densité (*source :* Wilkinson 1999, Hurter 2010)

Baudel (Baudel 2002 ; Baudel 2004) propose une méthode pour caractériser l'espace de visualisation et les variables utilisées sous la forme d'encapsulation d'objets. Il introduit le concept de **visualisation compacte**, qui se base sur le processus de visualisation (le pipeline) de Card et Mackinlay (Card et Mackinlay 1997). Toutefois, cette description est surtout utilisée pour caractériser le design de l'espace de visualisation, plutôt que les différentes étapes de projection de l'information. En ce sens, nous situons cette caractérisation dans la partie des « non RMP », car il n'y a pas de distinction nette entre les modèles de données et les modèles de la visualisation elle-même. La Figure 2-28 reprend un exemple dans la thèse de Hurter, issu des travaux de Baudel.

```
FillRectangle
{
    X = Get (Longitude)
    Y = Get (Latitude)
    Width = .04
    Height = .04
    FillRectangle
    {
       X =0 ; Y=0 ; Width = ½
       Height = Get (Crime)
    }
    FillRectangle
    {
       X = ½ ; Y=0 ; Width= ½
       Height = Get (Climate)
    }
}
```

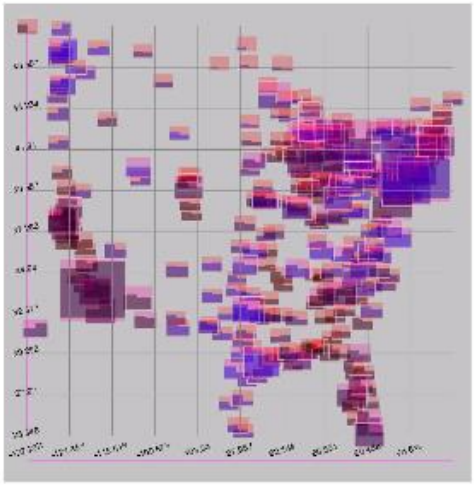


Figure 2-28. Caractérisation de Baudel pour la visualisation des statistiques de criminalité dans les grandes villes américaines densité (*source :* Hurter 2010)

Ces nombreux travaux de modélisation et de caractérisation montrent la diversité des approches possibles. Ces approches dépendent fortement de la finalité des modèles. Ainsi, la prochaine sous-partie présente l'utilisation de certains de ces modèles afin d'automatiser la visualisation ou d'étudier les visualisations adaptées à certains contextes.

2.2.3. Utilisation de formalisation pour une visualisation automatisée

Ces modélisations de méthodes de visualisation, lorsqu'elles sont adaptées à un langage machine, peuvent ensuite être utilisées dans des systèmes « intelligents » pour étudier et proposer des méthodes de visualisation adaptées aux données d'entrée. Ainsi, selon le modèle choisi, ces systèmes peuvent se baser sur la structure des données, leur typologie, ou les statistiques qui leurs sont associés.

En 1986, Mackinlay propose de développer un « langage graphique » pour caractériser **l'expressivité** (représentation de la bonne information) et **l'efficacité** (facilité à comprendre l'information véhiculée) de la représentation visuelle d'information. C'est ainsi qu'il développe le langage APT (*A Presentation Tool*) (Mackinlay 1986). Toutefois, ce langage construit des visualisations à la volée, ce qui engendre parfois des représentations inconnues et inhabituelles. Selon les caractéristiques des données étudiées, le langage APT permet de synthétiser un design graphique à partir de règles et de le tester par la suite.

Par la suite, les travaux de Mackinlay ont abouti au développement d'une nouvelle solution pour la visualisation d'information, avec le système *Show Me* (Mackinlay et al. 2007). Contrairement à APT, ce nouveau système ne crée pas de visualisations à la volée, mais permet de sélectionner une méthode prédéfinie et modélisée parmi une liste de possibilités. Ce nouveau système se base sur l'utilisation du langage VizQL pour créer des requêtes sur la structure des données (Stolte et al. 2002), d'après les règles de Bertin et l'algèbre d'APT. L'outil Show Me est aussi utilisé dans la solution commerciale *Tableau*[7]. La finalité de cet outil est de proposer des vues variées des mêmes données, que l'utilisateur peut sélectionner selon ses préférences. Néanmoins, cette approche se base encore sur le type de données prises en compte (N, O, Q, etc.).

De la même manière, le système *AutoVis* fait suite aux recherches menées par Wills et Wilkinson (Wills et Wilkinson 2010) afin de proposer une visualisation automatique de données choisies. AutoVis se base avant tout sur le processus d'analyse, tel qu'il serait effectivement mené par un opérateur humain lors de l'exploration de données. Les auteurs identifient les étapes successives suivantes :

- Reconnaissance du type de données en entrée du problème (source)
- Analyse de la structure des données
- Prioritiser les vues selon l'importance
- Résumer la structure des données en un affichage

Selon le type de données et la « grammaire des graphiques » (Wilkinson 1999), des visualisations adaptées sont alors proposées pour chacune de ces étapes, selon les propriétés intrinsèques des jeux

[7] www.tableausoftware.com

de données. Ainsi, cet outil permet d'accompagner l'utilisateur pendant l'exploration de données textuelles, relationnelles, images, etc., en proposant des outils adaptés. Dans cet article, les auteurs font la distinction entre un **système de visualisation automatique** et un **système de visualisation automatisé**. Un système automatique permet, selon eux, de décider de l'information la plus importante au sein de l'analyse de données, ainsi que de sélectionner l'information qui doit être visualisée. Un système automatisé, d'un autre côté, met automatiquement en forme les diagrammes et graphiques. Dans cette optique, AutoVis est un système de visualisation automatique, car il permet d'aider l'utilisateur dans son analyse. Au contraire, Show Me est un outil automatisé, qui selon les données et leur structure, met automatiquement en forme l'information visuelle.

Toutefois, ces nombreux travaux se basent essentiellement sur la théorie de la visualisation, telle que les variables de règles de Bertin, la grammaire des graphiques, ou encore les lois de la perception pré-attentive. La réalité des compétences et des habitudes de l'utilisateur n'entrent pas en compte dans ces propositions automatiques / automatisées de méthodes visualisation, et peuvent donc aboutir à des outils non adaptés à la situation de base. De plus, ces systèmes prennent rarement en compte une taxonomie des tâches à mener et des questions posées par l'utilisateur, desquelles dépendent fortement les visualisations à utiliser (Pfitzner et al. 2003).

D'autres travaux ont pu proposer des systèmes intelligents pour la visualisation d'information, en prenant en compte plus d'informations contextuelles sur l'utilisation effective de la visualisation. Gotz et Wen (Gotz et Wen 2009) ont basé leurs recherches sur le comportement effectif de l'utilisateur face à des outils de visualisation. Ils proposent alors des recommandations de visualisation, basées sur l'étude du comportement de l'utilisateur (*behavior-driven visualization recommendation*), à l'inverse des approches conduites par les données. Ce système étudie l'utilisation des méthodes de visualisation par les usagers, afin de détecter un motif d'utilisation (étape 1) ; ce motif permet ensuite de simuler l'analyse à venir et ainsi proposer des outils adaptés (étape 2). La Figure 2-29 résume l'architecture d'un tel système qui soutient dans l'analyse visuelle d'information.

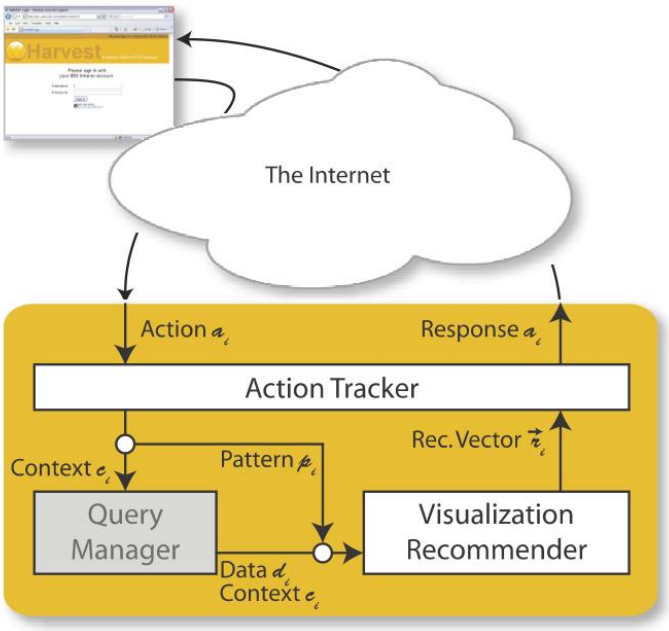


Figure 2-29. Architecture de la solution BDVR (*source :* Gotz et Wen 2009)

Comme nous l'avons vu, la philosophie « d'aide à la visualisation » a donc pu changer ces dernières années. La première préoccupation de la recherche a été de pouvoir proposer des visualisations qui soient adaptées aux types de données étudiées, ainsi qu'à leur structure propre. Cette façon de penser et d'étudier l'utilisation de visualisation permet en effet de voir l'information telle qu'elle est organisée, et d'en dégager des motifs prédominants, voire une logique dans son organisation. Par la suite, la recherche s'est portée sur une utilisation plus contextualisée, qui prend en compte le profil de l'utilisateur, ainsi que les tâches et questions qui le mènent à utiliser la visualisation. Ainsi, des méthodes ont été proposées pour étudier son utilisation effective de la visualisation, ses préférences et son mode d'exploration.

Toutefois, la recherche ne s'est pas portée sur l'utilisation de ces nombreuses connaissances dans un système formel qui puisse être partagé entre plusieurs utilisateurs, et être ainsi appliqué à de très diverses méthodes. Le langage ou le modèle utilisés peuvent être parfois compliqués à prendre en main et ne facilitent pas l'utilisation de ces systèmes à une échelle plus grande : plus de méthodes, plus de cas d'utilisation, etc.

La prochaine partie de ce chapitre présente donc une mode de formalisation des connaissances qui a pour but de partager une base commune de connaissances, tout en permettant l'application de règles pour un raisonnement automatisé. Cette technologie n'a pas été exploitée pour modéliser les nombreuses méthodes de visualisation, au sein d'un système de recommandation : nous présentons alors les possibilités et les axes qui nous ont intéressés pour notre recherche.

2.3. Formaliser des connaissances : recours aux ontologies

On retrouve le concept d'ontologie informatique dans de nombreux domaines, allant de l'ingénierie des connaissances à l'intelligence artificielle et le raisonnement. Aujourd'hui, son utilisation est surtout rapprochée à celle du web sémantique, afin de donner du sens au contenue des pages web. On rapproche l'ontologie de la représentation de connaissances, dont elle est une forme possible, et de l'interopérabilité des systèmes d'information.

Dans cette partie, nous présentons tout d'abord la définition d'une ontologie en informatique, ainsi que ses composantes et la manière de la construire. Nous présentons ensuite son rôle au sein des systèmes d'information, notamment pour l'information géographique.

2.3.1. Définition générale des ontologies

L'utilisation première du terme d'*ontologie* vient de la philosophie d'Aristote, et signifie dans ce contexte l'étude de l'être, de ses modalités et de ses propriétés. Aujourd'hui, de très nombreux domaines se sont appropriés ce concept et lui associent leur propre définition. En informatique, les ontologies sont liées au domaine de l'intelligence artificielle (IA) et l'utilisation de systèmes experts, notamment pour la formalisation de connaissances et pour le raisonnement sémantique.

La première définition informatique du concept d'ontologie est donnée par Gruber (Gruber 1993), comme étant une ***spécification explicite d'une conceptualisation***. Cette définition est ensuite précisée par Studer et al. (Studer et al. 1998) : une ***spécification formelle et explicite d'une conceptualisation partagée***. Les auteurs introduisent alors la notion de partage, qui se trouve être essentielle dans le rôle et l'utilisation des ontologies. Quatre termes clés sont retrouvés dans cette définition, qui donne son réel sens à l'ontologie : **formel** (standardisé, utilisable dans un langage

machine), **explicite** (définitions déclaratives), **conceptualisation** (abstraction du monde réel) et **partage** (utilisation de connaissances reconnues).

Une ontologie peut alors être représentée comme un graphe de concepts, traduisant tous des éléments du monde réel, et liés par des relations sémantiques (Staab et Studer 2009). Ce graphe est alors noté comme un quintuplet $O := \{C, R, H^C, rel, A^O\}$, avec :

- C et R sont des ensembles disjoints des concepts et des relations
- $H \subseteq C \times C$ est la hiérarchie des concepts
- $rel : R \rightarrow C \times C$ est une fonction spécifiant la relation entre les concepts
- A^O est un ensemble d'axiomes

S'appuyant sur une conceptualisation choisie, une ontologie ne peut pas représenter l'intégralité du domaine étudié. Il s'agit d'une représentation sélective du domaine, développée selon les besoins identifiés en termes d'informations et de requêtes possibles. La construction d'une ontologie n'est donc pas unique : elle est développée pour un objectif spécifique, qui rend alors son développement subjectif (Guarino 1998). Gruber (Gruber 1993) définit aussi plusieurs critères pour évaluer la qualité d'une ontologie :

- La clarté des concepts et sa documentation
- La cohérence, c'est-à-dire la non contradiction des axiomes, concepts et instances existants
- L'extensibilité, de manière à permettre l'ajout de nouveaux objets dans l'ontologie
- Une déformation d'encodage minimale, c'est-à-dire lorsque la spécification influe sur la conceptualisation
- Un engagement ontologique minimal, qui réduit l'ontologie à spécifier le plus simplement possible un domaine, une théorie. La base de connaissance, quant à elle, doit être plus complète et anticiper les requêtes possibles.

De plus, une ontologie doit respecter l'**hypothèse du monde ouvert** : ce qui n'est pas clairement dit comme faux, est considéré comme vrai. Ainsi, toutes les classes qui ne sont pas considérées comme distinctes sont automatiquement considérées comme ayant un recouvrement possible. Cette hypothèse est contraire au paradigme des bases de données et leur règle de « monde fermé » : si un objet n'a pas l'attribut recherché lors d'une requête, alors l'hypothèse testée est considérée comme fausse.

2.3.2. Constitution d'une ontologie

2.3.2.1. Composants d'une ontologie

Gruber (Gruber 1993 ; Gruber 2000) liste cinq types de composants nécessaires pour décrire une ontologie. Il s'agit des **concepts**, des **relations**, des **fonctions**, des **instances** et des **axiomes**. La Figure 2-30 représente une ontologie simple composée de cinq concepts (en bleu), deux individus (en orange) et six relations entre les concepts.

Les concepts, ou classes, sont la représentation abstraite d'objets matériels ou de notions qui sont décrits dans l'ontologie. Ce sont les concepts qui définissent toute la portée d'une ontologie et son niveau de détail. Lors de la création d'une ontologie, la difficulté est de savoir à quel niveau développer et spécifier ces concepts, de façon hiérarchique. Il est notamment défini par les liens qu'il

a avec les autres concepts de l'ontologie. Sur la Figure 2-30, les concepts représentés sont *Navire*, *Voilier*, *Bateau à moteur*, *Voile* et *Moteur*. Cette représentation simple de la connaissance ne descend pas vers des détails plus fins et se contente de quelques liens pour définir ces concepts.

Les relations sont les liens qui existent entre les différents concepts de l'ontologie. Elles peuvent être des liens sémantique, ou bien des liens de subsomption (hiérarchiques) tels que les liens de parenté *est-un*. Une relation sémantique est définie par le sens, ou contenu sémantique, du lien entre les concepts qu'elle relie (Bachimont 2000). Par exemple, sur la Figure 2-30, les liens hiérarchiques *est-un* relient les concepts *Voilier* et *Bateau à moteur* à *Navire*, qui est le concept parent. *Se-propulse-par* est alors une autre relation entre concepts, qui définit les spécificités des concepts *Voilier* et *Bateau à moteur*. Ces relations peuvent être, ou non, transitives ou symétriques.

Les fonctions sont des types spécifiques de relations, où un des éléments est défini selon les éléments précédents.

Les instances, ou individus, sont les individus qui peuplent une ontologie. Il s'agit d'une instanciation des concepts, avec des valeurs spécifiques pour leurs attributs. En peuplant ainsi une ontologie, on la fait évoluer d'une modélisation vers une réelle **base de connaissances**.

Les axiomes sont des assertions vraies, associées à des concepts ou des instances. Ils peuvent être utilisés pour définir des règles de raisonnement au sein de l'ontologie. Leur utilisation permet, entre autres, de vérifier la validité des informations, c'est-à-dire la cohérence de l'ontologie.

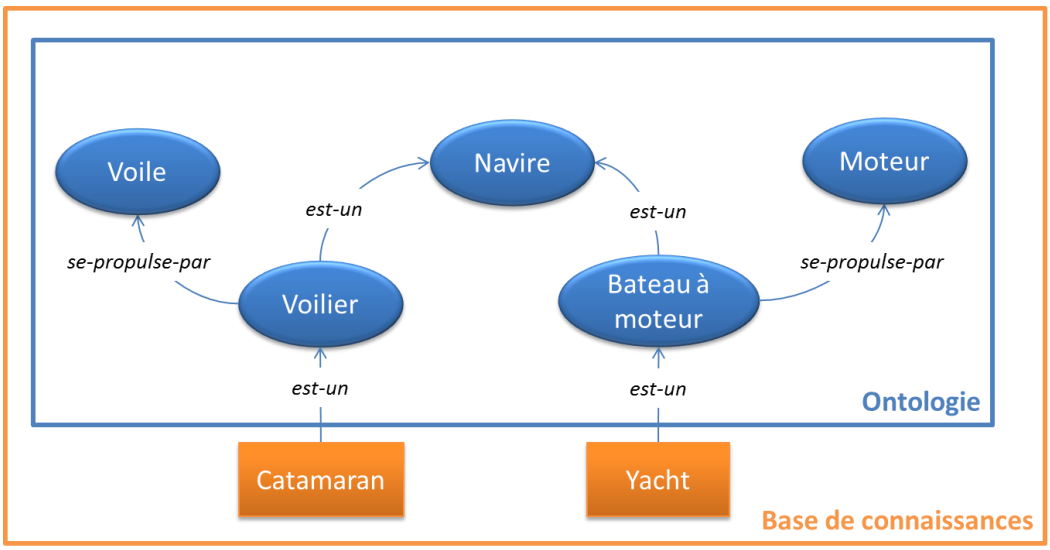


Figure 2-30. Base de connaissances à quatre concepts (bleu) et deux individus (orange)

La structure d'une base de connaissances (Staab et Studer 2009) est donc de la forme $KB = \{O, I, inst, instr\}$ avec :

- O une ontologie, telle que définie plus haut
- I est un ensemble d'individus
- $inst : C \rightarrow 2^I$ est la fonction d'instanciation de concepts
- $instr : C \rightarrow 2^{I \times I}$ est la fonction d'instanciation de relations

Plusieurs niveaux d'ontologies peuvent donc être distingués, qui vont avoir des finalités différentes (Hernandez et Mothe 2006 ; Heer et al. 2009) :

- une ontologie *légère*, qui ne prend en compte que des concepts et relations
- une ontologie *lourde*, qui introduit des attributs et des fonctions
- une *base de connaissances*, grâce aux règles et instances qui sont ajoutées

De très nombreux ouvrages et thèses se sont intéressés à la théorie des ontologies informatiques, leur mode de construction et leur fonctionnement pour la gestion de l'information et des connaissances. La thèse de Vandecasteele (Vandecasteele 2012) mène un état de l'art très complet sur cette partie théorique et sur l'intégration de la dimension spatiale et topologique au sein des ontologies. L'ouvrage par Staab et Studer (Staab et Studer 2009) décrit de manière très précise la théorie des ontologies et des liens entre concepts.

Dans la sous-partie suivante, nous présentons le raisonnement au sein d'une ontologie, à partir des concepts développés, grâce à l'utilisation de règles.

2.3.2.2. Langages utilisés

Des langages standardisés permettent de développer des ontologies, notamment RDF et OWL (Bechhofer 2002 ; Bechhofer et al. 2002). Le langage RDF (*Ressource Description Framework*) permet d'annoter les ressources du web, de manière à être interprétables par la machine (web sémantique). Ce langage permet de décrire des graphes, reliant des objets par des prédicats. En reprenant l'exemple de la Figure 2-30, la représentation en langage RDF de l'instance *Voilier* du concept *Navire* serait alors :

```
<rdf:RDF
   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
   xmlns:dcterms="http://purl.org/dc/terms/">
      <rdf:Description rdf:about= "Voilier">
         <rdf:is-a>Navire</rdf:is-a>
      </rdf:Description>
</rdf:RDF>
```

Toutefois, le RDF a été concu de manière à être le plus générique possible : cela pose alors des limites importantes pour la description de situations complexes (Antoniou et van Harmelen 2009) : cardinalité, inclusion, exclusion, etc. Afin de faire face à ces limites, le RDF a ensuite été étendu grâce au langage OWL (*Ontology Web Language*). Son utilisation dans l'architecture du web sémantique est schématisé comme le pivot au sein du « millefeuille » de la Figure 2-31. Ce langage permet les requêtes, les règles et la comparaison des classes. De manière à simplifier son utilisation, plusieurs versions ont été proposées : OWL-Lite, OWL DL et OWL Full. Le langage OWL a ensuite été amélioré dans une nouvelle version révisée, OWL 2 (Grau et al. 2008). Ainsi, de nouveaux axiomes y sont ajoutés, selon les besoins de la communauté scientifique (Patel-Schneider et Horrocks 2006).

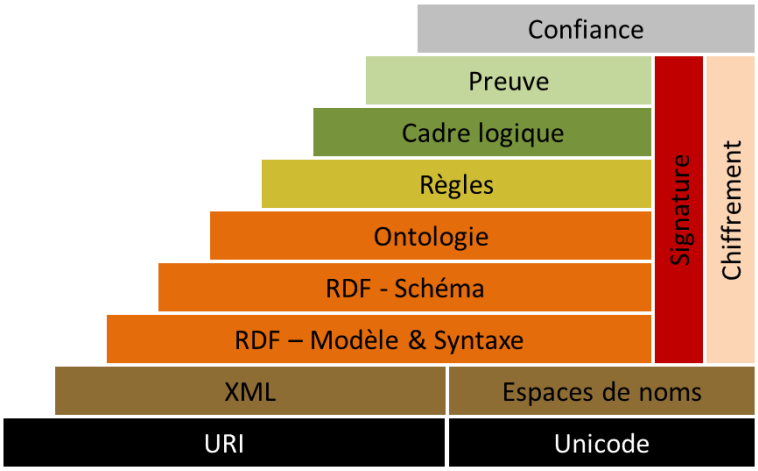


Figure 2-31. Architecture du web sémantique, selon le W3C (d'après les propositions du W3C)

2.3.2.3. Développement d'une ontologie : l'approche hybride

Plusieurs approches permettent d'identifier et développer les concepts d'une ontologie. Ces approches peuvent être de type top-down (descendante), ou bien bottom-up (ascendante). La première catégorie permet de partir de l'analyse des sources d'informations afin d'en extraire des concepts d'intérêt (López-Pellicer et al. 2007) ; tandis que l'approche bottom-up part des objets considérés comme « indivisibles », ou *primitives*, afin de remonter vers les objets et règles qu'ils peuvent décrire (van der Vet et Mars 1998). L'approche bottom-up part donc de primitives qui sont souvent abstraites, afin de décrire des objets complets et concrets. Suivant cette logique, on parle souvent d'ontologie de haut niveau et d'ontologie de bas niveau : cela revient donc à considérer respectivement les concepts de ces ontologies comme très généraux, ou bien comme très spécifique (le plus bas niveau de l'ontologie, les primitives). Lors du développement d'une ontologie, l'une des grandes difficultés est d'identifier le niveau de primitive acceptable : par définition, une ontologie est un choix de conceptualisation, et ne peut donc pas décrire l'intégralité de la structure des données.

En plus de ces méthodologies, l'approche hybride (López-Pellicer et al. 2007) pour le développement d'une ontologie combine à la fois une approche top-down et bottom-up que nous allons décrire ici. La Figure 2-32 représente les différentes étapes de ce processus qui aboutit à une ontologie du domaine choisi.

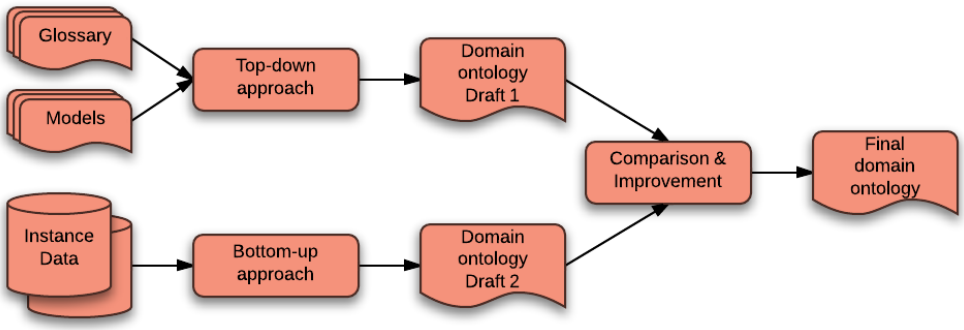


Figure 2-32. Approche hybride pour le développement d'une ontologie (d'après López-Pellicer et al. 2007)

Les approches top-down et bottom-up s'effectuent de manière parallèle et complémentaire, afin de comparer, compléter et améliorer les concepts et relations identifiés lors de ces procédés de modélisation. On reconnait l'approche descendante qui part des modèles généraux et d'un glossaire

pour un domaine choisi, afin d'aboutir à une première version d'ontologie. L'approche ascendante part, quant à elle, des données concrètes qui sont disponibles, afin de pouvoir remonter vers une conceptualisation général dans une seconde ontologie.

2.3.2.3.1. Approche descendante

L'approche hybride proposée par López-Pellicer et al. (López-Pellicer et al. 2007) s'appuie sur la méthode descendante Methontology (Gómez-Pérez et al. 2004) décrite sur la Figure 2-33. Cette méthodologie permet, en partant d'un glossaire de terme du domaine choisi, d'aboutir à un ensemble de concepts, relations et axiomes qui forment l'ontologie. Cette méthode peut être appliquée aux domaines distincts que sont l'analyse géovisuelle et l'analyse des comportements des navires, dans l'optique de gestion des risques maritimes. Chacun de ces domaines a en effet son propre vocabulaire et ses propres règles qui le décrivent. Pour le domaine maritime, nous utilisons surtout cette approche afin d'identifier les concepts pour la caractérisation des comportements et des risques : ce sont les concepts de plus haut niveau, à l'échelle de notre étude. La méthodologie Methontology se compose de trois étapes successives (Gómez-Pérez et al. 2004 ; Bendaoud et al. 2007) :

- L'extraction de termes : cela peut être fait par une fouille de données automatisée ou non, par exemple. Cette étape permet de mettre au point une taxonomie des concepts. Des propriétés et relations sont aussi identifiées lors de cette étape.
- La construction du noyau de l'ontologie : les termes et les relations identifiées précédemment définissent une hiérarchie.
- L'extraction de relations transversales, qui permet ensuite de relier des éléments de niveaux différents, en complément de la hiérarchie simple des taxons.

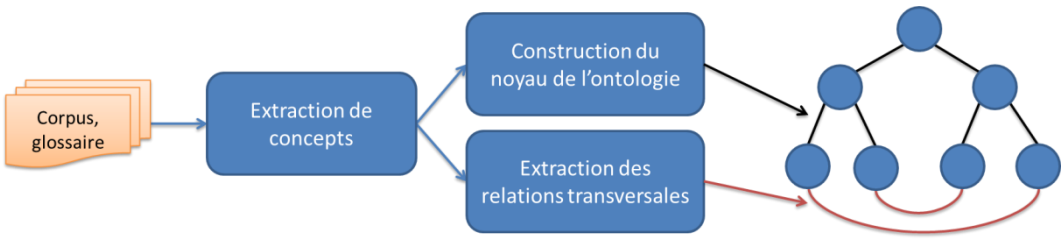


Figure 2-33. Approche descendante, pour la conceptualisation des éléments de bas niveau, par la méthode Methontology (d'après Bendaoud et al. 2007)

2.3.2.3.2. Approche ascendante

En suivant l'approche proposée par López-Pellicer et al. (López-Pellicer et al. 2007), la seconde étape de modélisation est l'approche ascendante : cette étape se base donc sur les données détaillées et concrètes du domaine étudié. Dans notre cas, de telles données peuvent être les informations AIS et les données qui les composent (positions des navires, vitesse, etc.) pour le domaine maritime, ou bien les variables visuelles et dynamique pour le domaine de la visualisation. De ces informations de très bas niveaux, qui sont considérées comme les primitives de l'ontologie, des regroupements successifs sont opérés afin de créer des généralisations d'ordre sémantique. La Figure 2-34 représente une stratégie d'approche ascendante pour la constitution d'une ontologie : cette méthode présentée est appelée *analyse de concepts formels*, ou *Formal Concept Analysis* (FCA) en anglais (Wille 1982 ; Ganter et Wille 1999). L'analyse de concepts formels repose sur trois étapes successives :

- L'extraction du contexte à partir d'un corpus de données (objets)
- L'analyse des concepts, par la recherche de propriétés partagées entre les objets
- La création de l'ontologie par la mise en relation de ces concepts

Une fois ces deux approches appliquées, les concepts de plus haut et de plus bas niveau des ontologies résultantes peuvent alors être comparés et améliorés, de façon à aboutir à une structure commune de ces concepts.

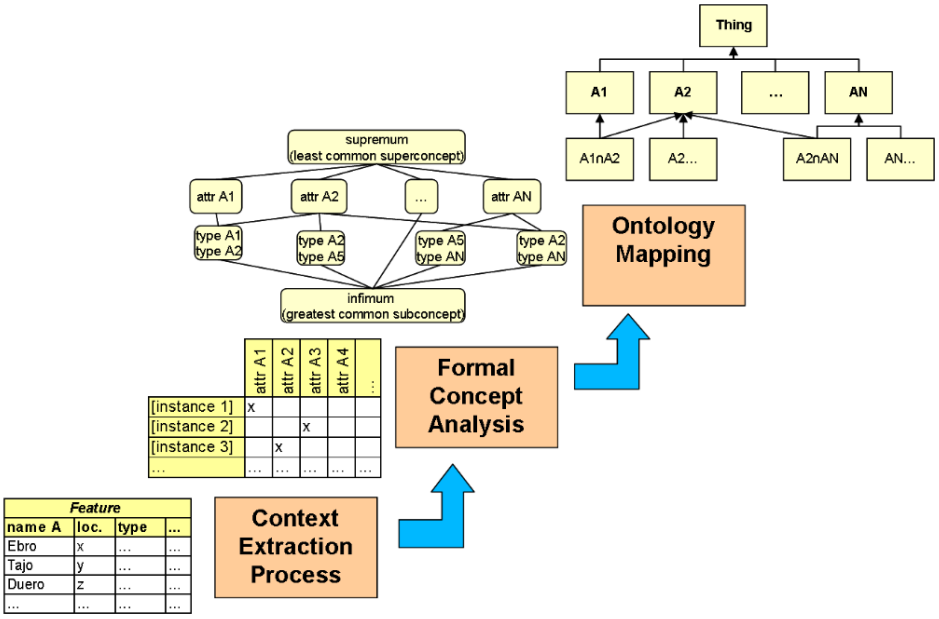


Figure 2-34. Approche ascendante, pour la conceptualisation des éléments de haut niveau (*source :* López-Pellicer et al. 2007)

2.3.3. Raisonnement : utilisation de règles et inférence

Le développement de l'ontologie, puis de la base de connaissances par l'ajout d'instances, permet de définir et modéliser un domaine spécifique. Une fois cette étape achevée, des règles de raisonnement peuvent être ajoutées afin d'analyser le contenue de la base de connaissances. Celles-ci ont pour seule utilisation l'ajout de nouvelles relations, c'est-à-dire l'application de relations définies sur des concepts ou des instances qui existent déjà dans l'ontologie. Les règles ne peuvent pas mettre à jour des relations, ou créer de nouveaux objets. Le langage utilisé est le *Semantic Web Rule Langage* (SWRL). Dans cette sous-partie, nous présentons l'utilisation du SWRL, son formalisme et l'application des règles par inférence.

2.3.3.1. Règles SWRL

Le SWRL provient des langages OWL-DL pour la description de l'ontologie, et RuleML (*Rule Markup Language*) pour la formalisation des règles (Horrocks et al. 2004). Les règles sont de la forme *antécédents* → *conséquences*, formalisées de la manière suivante :

$$a(x,y) \wedge b(y,z) \wedge c(x) \ldots \rightarrow n(x,y)$$

- a, b et n sont des prédicats binaires tels que des relations
- c est un prédicat unaire, tel qu'un concept

- x et y sont des variables, telles que des nombres ou objets de l'ontologie

Ces règles sont donc définies à partir de concepts et d'instances déjà existants au sein de l'ontologie, il en est de même pour la conséquence de cette règle. Cette formulation s'appuie sur les prédicats binaires et unaires OWL, mais peut aussi prendre en compte des fonctions mathématiques, de logique booléenne, d'opérations sur les chaines, dates, etc., grâce aux *buit-in* autorisés (Horrocks et al. 2004 ; Grau et al. 2008). Des comparaisons entre les objets et les valeurs de leurs attributs sont donc autorisés dans les règles SWRL. Le formalisme SWRL prend alors la forme suivante, lors de l'écriture des règles :

$$ConceptA(?\,var1) \wedge ConceptB(?\,var2) \wedge Relation1(?\,var1, ?\,var2) \rightarrow Relation2(?\,var2, ?\,var1)$$

Sur cet exemple, *ConceptA* et *ConceptB* sont des concepts de l'ontologie, *Relation1* et *Relation2* sont des relations définies. *?var1* et *?var2* sont alors les notations de variables, qui doivent être respectivement, sur cette règle, être de type *ConceptA* et *ConceptB*.

Afin d'appliquer les règles définies dans une ontologie, il est nécessaire d'utiliser un moteur d'inférence qui va parcourir la liste de règles.

2.3.3.2. Processus d'inférence

De très nombreux raisonneurs existent et supportent le langage SWRL. Plusieurs d'entre eux sont automatiquement intégrés aux logiciels de gestion d'ontologies, tels que FaCT, FaCT++ et Hermit. D'autres, plus complexes, doivent être ajoutés : c'est le cas par exemple du raisonneur Pellet.

Le moteur d'inférence, ou moteur de raisonnement, a pour rôle d'inférer de nouveaux liens entre les concepts et instances, en appliquant les axiomes de l'ontologie. Celui-ci va aussi permettre de découvrir de possibles erreur de cohérence au sein de l'ontologie : celles-ci peuvent provenir de la structure de l'ontologie, des connaissances, ou encore de règles contradictoires.

La

Table 2-6 présente une comparaison de plusieurs moteurs d'inférences, dans l'article par Dalwadi et al. (Dalwadi et al. 2012). Dans cette étude, les auteurs comparent notamment le mode de fonctionnement du raisonneur (chaînage et logique), la vérification de cohérence, le langage de requêtes et le langage de règles supportés. Nous reviendrons sur ce tableau et les différentes solutions proposées, lors du Chapitre 3 de cette thèse.

Table 2-6. Comparaison de moteurs d'inférence (d'après Dalwadi et al. 2012)

	FaCT	FaCT++	Pellet	Racer	Hoolet	F-OWL	Hermit	KaON2
Licence	Libre (open source)	Libre (open source)	Libre (open source)	Non libre	Libre (open source)	Libre (open source)	Libre (open source)	Libre (open source)
Plateforme	Windows / Linux	Windows / Linux	Windows	Windows	Linux	Windows	Windows	Windows / Linux
Support OWL	OWL-DL	OWL-DL	OWL-DL	OWL-DL	OWL-DL	OWL-Full	OWL-DL	OWL-DL
Algorithme raisonnement	Tableau	Tableau	Tableau	Tableau	Preuve 1er ordre	Tableau	Hyper-Tableau	Calculs de résolution 1er ordre
Test cohérence	✓	✓	✓	✓	-	✓	✓	✓
Interface	DIG, commandes	DIG, commandes	DIG, JAVA	DIG, JAVA, GUI	JAVA	JAVA, GUI, commandes	Commandes	JAVA, GUI, commandes
Langage requête	-	-	RDQL, SPARQL	Racer Query Language	-	Frame Style, RDQL	-	SPARQL
Langage règles	-	-	SWRL – DL safe	SWRL	SWRL	SWRL	SWRL – DL safe	SWRL – DL safe

Conclusion du Chapitre 2

Dans ce chapitre, nous avons commencé par présenter les travaux relatifs à la visualisation d'information, afin de présenter le domaine de l'analyse géovisuelle. Ce domaine propose de très nombreuses méthodes visuelles pour représenter l'information, qui comporte une composante spatiale, afin de supporter notre exploration et notre analyse des données. Pour cela, la visualisation se repose sur notre perception de l'information par des regroupements automatiques des données, ainsi qu'une perception plus contrôlée par du texte ou des images, par exemple. Ces solutions ont généralement un haut niveau d'interaction, pour nous permettre de manipuler l'information visualisée et ainsi créer des sous-ensembles de données pour nous aider à mieux comprendre les possibles motifs, par la visualisation. La visualisation d'information géographique, et a fortiori l'analyse géovisuelle, jouent un rôle majeur dans la compréhension de situations (*Situation Awareness*) et la prise de décisions qui en découle. Néanmoins, on ne peut pas les qualifier ni de PSE, car ces méthodes ne fournissent pas une réponse, ni de SAD car elles ne sont pas fondées sur des connaissances formalisées. La visualisation est donc un outil qui aide pour la prise décision, mais n'est pas un SAD tel que la définition originale les présentent. De plus, un problème majeur subsiste encore dans l'utilisation de la visualisation d'information géographique, ainsi que le renseignement de leurs métadonnées. Il n'existe pas de catalogue de ces très nombreuses méthodes, ni de support dans leur utilisation. De nombreux environnements, développés par les centres de recherche, proposent des listes de fonctionnalités importantes et sont généralement indépendant du domaine d'application. Toutefois, chaque environnement ou logiciel comporte ses propres outils spécialisés, ou encore une architecture et un langage très spécifiques. Nous pouvons citer, parmi les plus importants dans le domaine des sciences de l'information géographique :

- Map4Decision[8] est une solution web pour les applications SOLAP et l'étude du territoire (Bédard et al. 2007 ; Bédard 2010), développée par le CRG de l'Université de Laval (Québec)
- GeoVISTA Studio[9], une plateforme pour le traitement, l'analyse et la visualisation de données géographiques (Takatsuka et Gahegan 2002 ; MacEachren et al. 2004). Elle est issue des travaux du GeoVISTA Center, Penn State University (USA).
- ILWIS[10] est le résultat de nombreux travaux d'analyse spatiale et de télédétection (Meijerink et al. 1988), développé à l'ITC, University of Twente (Pays-Bas). Depuis quelques années, c'est le consortium 52° North qui est en charge de la maintenance de ce logiciel.
- CommonGIS[11] est une solution logicielle commerciale, pour la cartographie interactive et la visualisation avancée de données spatiales (Andrienko et Andrienko 1999 ; Andrienko et al. 2002). Cet outil est le fruit de plusieurs années de R&D au centre IAIS, Fraunhofer Institute (Allemagne).

Contrairement aux SIG, tels qu'ArcGIS, MapInfo ou QGIS, ces solutions sont beaucoup plus adaptées à l'exploration et à l'analyse interactive de données spatiales, sémantiques et temporelles. Ces

[8] http://www.intelli3.com/map4decision
[9] http://www.geovistastudio.psu.edu
[10] http://www.ilwis.org/
[11] http://commongis.com/

fonctionnalités avancées pour l'exploration de données aux structures encore inconnues mettent en évidence la différence fondamentale entre l'utilisation de SIG et de la géovisualisation – et a fortiori de l'analyse géovisuelle (MacEachren 1995 ; MacEachren et Kraak 1997). Toutefois, l'utilisation de ces outils reste encore limitée lorsqu'un utilisateur non expert en information géographique, ou non formé aux logiciels, doit analyser des informations, extraire des connaissances et prendre des décisions. Une formation conséquente doit alors être suivie afin de comprendre les possibilités de tels logiciels et des visualisations qu'ils proposent. A moins d'en être le concepteur, il est souvent difficile de savoir quel outil utiliser, quels paramètres choisir et comment interpréter les résultats de la visualisation résultante.

En analyse géovisuelle, il n'existe pas de support ou de guide à l'utilisation, lorsque l'utilisateur doit définir un processus d'analyse géovisuelle, choisir les méthodes de visualisation en conséquence et sélectionner les données appropriées pour son étude. Si nous revenons sur le modèle TAM pour l'acceptation de la technologie (Davis et al. 1989 ; Davis 1993), les nœuds centraux pour jouer sur l'acceptation et l'utilisation effective d'une technologie, telle que les méthodes d'analyse géovisuelle, sont l'utilité perçue et la facilité d'utilisation perçue. Afin que les environnements d'analyse géovisuelle puissent être effectivement acceptés et utilisés par des opérationnels, il est donc nécessaire que l'utilisateur **comprenne** leur **utilité** pour son propre travail, ainsi que leur **mode d'utilisation** pour l'analyse de données.

Dans la seconde partie de ce chapitre, nous avons donc présenté les différents travaux relatifs à la caractérisation et la modélisation de méthodes de visualisation. Chaque solution peut en fait être assimilée à une transformation de la donnée brute par une projection vers un espace visuel, tout en étant contrôlée à chaque étape par l'utilisateur. De nombreux travaux ont pu proposer des systèmes de visualisation automatiques ou automatisés, afin de visualiser au mieux l'information la plus importante. Néanmoins, toutes ces études se sont basées sur d'importants travaux sur des langages graphiques, qui sont rarement interopérables. Cela pose des problèmes d'échange d'information, entre humains ou entre systèmes. Nous nous sommes donc intéressés à un mode spécifique de la formalisation des connaissances, qui puisse être compris et partagé par tous : les ontologies.

Nous avons vu en dernière partie la définition générale des ontologies et de ses composantes. Une ontologie est un modèle biaisé, car issu des choix relatifs à la conceptualisation du domaine choisi. Celle-ci doit être formelle et précise, selon le niveau d'information et le type de requête que l'on veut effectuer. Mais une ontologie reste un modèle qui peut être partagé, être mis à jour et évoluer grâce aux travaux des communautés concernées. Pour cela, nous avons choisi d'utiliser ce type de formalisation pour décrire les solutions de visualisation, pour une meilleure intégration et utilisation au sein de systèmes d'information tels que les systèmes de surveillance de trafic.

CHAPITRE 3 - CONCEPTION D'UN ENVIRONNEMENT D'AIDE A L'ANALYSE GEOVISUELLE POUR LES RISQUES MARITIMES

Dans les deux premiers chapitres de ce manuscrit de thèse, nous avons présenté les travaux existants dans le domaine de l'analyse géovisuelle, notamment pour l'exploration et l'analyse de données de mouvement. Les grandes limites de ces travaux se retrouvent notamment dans la « non utilisation opérationnelle » de ces technologies de l'information et le manque de métadonnées sur leur utilisabilité. Pour cela, nous avons présenté les travaux menés dans le domaine de la modélisation de la visualisation d'information, ainsi que des systèmes intelligents pour la visualisation automatique et / ou automatisé d'information. Toutefois, ces travaux n'ont pas été appliqués au domaine de l'information géographique. Nous pouvons observer aujourd'hui de nombreux catalogues de données et de métadonnées afin de caractériser et normaliser l'information géographique (infrastructures de données géographiques, INSPIRE, etc.). Toutefois, il n'existe pas de d'équivalent appliqué aux méthodes, tel qu'une catégorisation et un catalogage des méthodes de visualisation d'information géographique.

Dans ce chapitre, nous présentons notre travail de conception d'un environnement d'aide à l'analyse géovisuelle, pour l'étude des mouvements d'objets. Nous appliquons cette recherche au domaine de la sécurité et sûreté de la maritimisation de l'énergie, notamment à l'analyse des mouvements de navires.

La première partie de ce chapitre présente la méthodologie proposée pour la définition et la conception d'un système d'aide à l'analyse géovisuelle. Nous revenons notamment sur l'utilisation d'ontologie pour la formalisation de la connaissance, qui se base sur des concepts, des instances et des règles. Ainsi, dans la deuxième partie, nous présentons les concepts que nous manipulons au sein de notre base de connaissances, puis nous présentons le modèle ontologique complet. Enfin, dans une troisième partie, nous expliquons l'utilisation de règles qui se fondent sur cette base de connaissances, qui servent de moteurs pour le raisonnement ontologique.

3.1. Méthodologie proposée

Cette première partie du chapitre présente la méthodologie pour le travail de formalisation que nous avons adoptée durant notre recherche. Nous commençons par définir ce que nous appelons un *environnement d'aide à l'analyse géovisuelle*, qui est un nouveau concept au cœur de cette thèse. Par la suite, nous présentons nos choix pour développer un tel environnement, ainsi que la méthodologie de développement des objets qui lui sont associés.

3.1.1. Un environnement d'aide à l'analyse géovisuelle

Comme nous l'avons vu précédemment, les logiciels et environnements permettant l'analyse géovisuelle, tels que GeoVista, CommonGIS, ILWIS, etc., ne proposent pas de classification claire et efficace des nombreux outils de visualisation d'information géographique qu'ils comportent. Bien que ces outils apportent une contribution importante au processus d'analyse spatio-temporelle, il peut s'avérer compliqué de savoir dans quel cas les utiliser ou non. Pour cela, il est essentiel d'étudier une stratégie afin de décrire l'intérêt de chaque méthode de visualisation, et ainsi stimuler leur utilisation. C'est dans ce cadre que nous proposons un nouveau terme, qui fait suite à la philosophie des systèmes d'aide à la décision et environnement de résolution de problèmes, mais dont le rôle se situerait en amont de ceux-ci. Nous introduisons donc le terme d' ***« environnement d'aide à l'analyse géovisuelle »*** (EAAG), ou ***Geovisual analytics support environment*** en anglais.

La prochaine sous-partie définit en détail ce terme et son utilisation, qui s'inscrit donc au sein du processus d'analyse des comportements maritimes à risques et de la prise de décision par la visualisation d'information.

3.1.1.1. Proposition d'une définition

Le but d'un tel environnement est d'assister les opérateurs lors de l'analyse de données, par la visualisation, en utilisant des méthodes adéquates selon leur cas d'utilisation. Nous proposons donc une définition de cet environnement, qui puisse aider dans l'utilisation de méthodes de visualisation pour l'exploration et l'analyse d'information géographique. En suivant la définition des environnements d'aide à la décision, notre définition d'un environnement d'aide à l'analyse géovisuelle est la suivante.

Un environnement d'aide à l'analyse géovisuelle est un système informatique qui permet de guider un utilisateur dans son utilisation des méthodes d'analyse géovisuelle. Afin de résoudre une classe de problèmes, telle que l'analyse d'informations géographiques choisies, un EAAG propose des solutions adaptées au cas d'utilisation. Les problèmes posés peuvent être dépendants d'un domaine qui utilise un vocabulaire spécifique, tel que le contrôle de trafic aérien, le contrôle du trafic maritime, etc. Les problèmes à résoudre peuvent aussi partager un sens générique, et être ainsi retrouvés dans plusieurs domaines. L'exploration et l'analyse des trajectoires d'objets mobiles, par exemple, obéit à des stratégies similaires selon les domaines d'application, et dispose donc d'un vocabulaire générique, tel que l'on a pu le voir dans la partie 1.3.2 de cette thèse.

Contrairement aux SAD et aux PSE, les EAAG n'ont pas pour finalité de proposer des solutions à des problèmes, mais de faciliter l'utilisation de méthodes adéquates pour résoudre ces problèmes. De façon similaire aux PSE et aux SAD, ils doivent s'adapter aux différents contextes d'utilisation de manière à proposer des solutions adaptées au cas d'utilisation. Un choix doit néanmoins être laissé à l'utilisateur, de manière à ce qu'il puisse lui-même décider de la méthode, ou des méthodes, qu'il préfère utiliser. Le processus de prise de décision est alors entièrement laissé à l'être humain, et non pas à la machine.

Tel que nous l'avons vu précédemment, la flexibilité d'un tel environnement doit être permise par son développement formel, c'est-à-dire interprétable par la machine, et générique. Un EAAG doit alors présenter les caractéristiques suivantes :

- Inclure un catalogue varié de méthodes de visualisation d'information à composantes géographiques, qui pourront être proposées selon les nombreux cas d'utilisation.
- Proposer plusieurs solutions de visualisations, qui puissent être utilisées de manière isolée, ou bien liées ensemble au sein d'un environnement d'analyse géovisuelle.
- Un EAAG doit être distinct des méthodes d'analyse géovisuelle qu'il propose : c'est un système indépendant du processus et des méthodes de visualisation, et qui ne doit pas dépendre des technologies utilisées par les visualisations proposées (web, logiciel, etc.).
- Un choix et un accès rapide à la méthode ou aux méthodes d'analyse géovisuelle doit être possible. De cette façon, le temps de réflexion pour la stratégie de visualisation et pour le choix des méthodes peut être réduit. Le temps est ensuite majoritairement utilisé pour l'analyse de l'information en elle-même.

La Figure 3-1 résume la situation et le mode d'utilisation d'un environnement d'aide à l'analyse géovisuelle, qu'il faut donc replacer dans le processus d'analyse de l'information et de prise de décision. La Figure 3-2 montre alors sa place à une échelle plus large dans ce processus, basé sur la boucle du modèle COCOM pour le contrôle d'une situation.

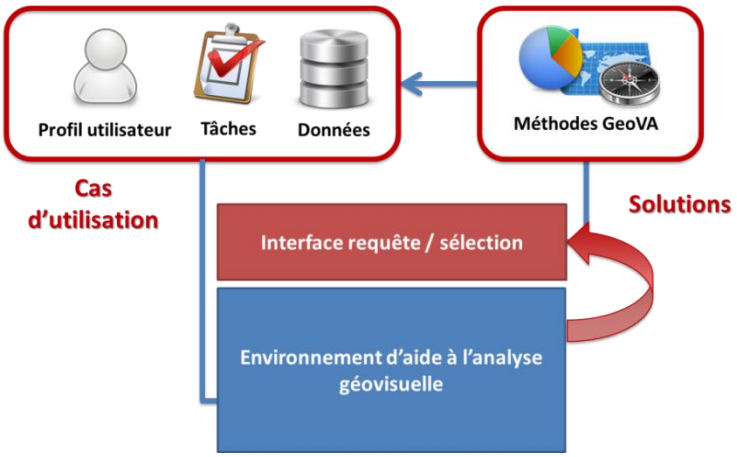


Figure 3-1. Mode d'utilisation d'un environnement d'aide à l'analyse géovisuelle

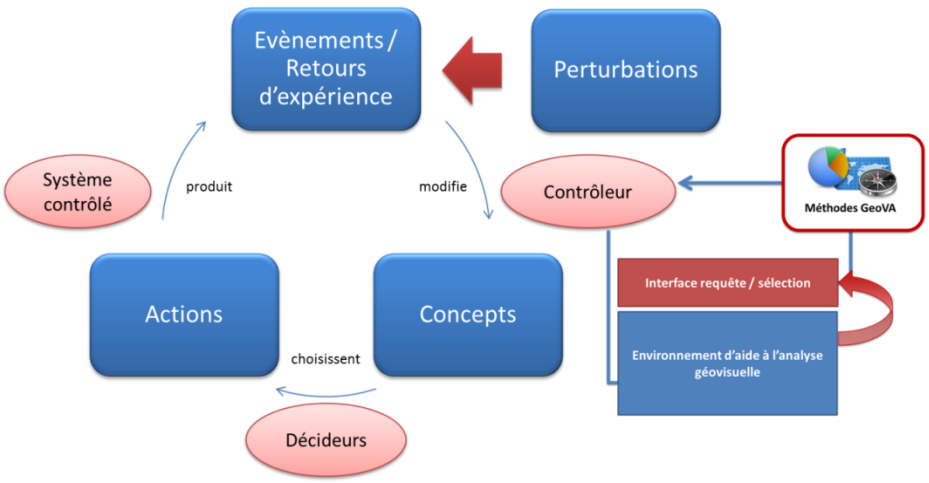


Figure 3-2. Place d'un EAAG au sein de la prise de décision, intégré dans le modèle COCOM

Après avoir donné notre définition d'un EAAG, nous proposons maintenant d'expliquer le développement d'un tel environnement. Dans la sous-partie suivante, nous présentons donc un modèle général pour la réalisation d'un EAAG, ainsi que les concepts qui doivent alors être pris en compte durant ce processus.

3.1.1.2. Les éléments d'un EEAG

Nous l'avons vu précédemment, le choix d'une méthode pour visualiser, explorer et analyser des informations dépend de très nombreux paramètres. Tout comme les PSE et SAD, le contexte d'utilisation est notamment très important pour une utilisation adaptée de la visualisation d'informations géographiques. Ce problème et les prémices d'une solution ont été proposés lors de première conférences, au cours de notre recherche (Vatin et Napoli 2013a ; Vatin et Napoli 2013c). On retrouve les informations fondamentales pour caractériser ce contexte dans des travaux similaires :

- Le **profil** de l'utilisateur, caractérisé par de nombreuses spécificités (Slocum et al. 2001 ; Endsley et al. 2003 ; Mountain 2005 ; Harding 2013). L'opérateur ou l'analyste vont avoir des rôles et des caractéristiques différentes selon leur formation, par exemple.
- Les **données** qui sont disponibles lors de cette analyse, ainsi que leur typologie (Mackinlay et al. 2007 ; Wills et Wilkinson 2010).
- L'**environnement** dans lequel l'étude des données est faite, notamment le milieu géographique et les ressources technologiques. L'environnement peut être mobile, comme lors de l'utilisation d'un système embarqué (Mallé-Noyon 2008 ; Lopez-Velasco et al. 2009).
- Le **but d'utilisation** de l'analyse géovisuelle, c'est-à-dire les questions posées par l'utilisateur pendant son utilisation. De façon à répondre un but spécifique, un ensemble de tâches successives forme alors une stratégie d'analyse visuelle de l'information (Fabrikant 2001 ; Andrienko et al. 2005 ; Gotz et Zhou 2008 ; Andrienko et Andrienko 2013).

L'autre aspect à prendre en compte est donc le cœur de cette recherche : les solutions d'analyse géovisuelle qui peuvent être utilisées dans l'analyse des données. Pour cela, il est nécessaire de connaitre les possibilités offertes par chacune de ces méthodes, puis de les comparer aux éléments identifiés précédemment. Par exemple, il est nécessaire de connaitre les données qui peuvent être visualisées (selon leur nombre, leur type, etc.), les interactions possibles avec l'information, les tâches d'analyses autorisées, les limites d'utilisation, les dimensions de la visualisation, etc.

Dans la prochaine partie, nous présentons donc nos choix techniques pour la proposition d'un environnement d'aide à l'analyse géovisuelle. Pour cela, nous avons exploité le concept de base de connaissances, tel qu'il a pu être présenté dans le chapitre 2, partie 2.3.

3.1.2. Un système à base de connaissances pour l'analyse géovisuelle appliquée au domaine maritime

Dans la sous-partie précédente, nous avons expliqué ce que nous appelons un environnement d'aide à l'analyse géovisuelle. De manière à pouvoir s'adapter aux domaines d'application et aux cas d'utilisation, un EAAG doit donc être fondé sur une représentation formelle de la connaissance : cette représentation doit être claire, partageable et interprétable par un langage machine. Pour cela, nous présentons dans cette sous-partie les choix techniques que nous avons effectué pour le

développement d'un système à base de connaissances, qui puisse représenter le domaine de l'analyse géovisuelle et les besoins en sûreté et sécurité de la maritimisation de l'énergie.

Dans les sous-parties qui suivent, nous présentons donc les différentes connaissances à formaliser, le rôle des règles au sein de cette formalisation, et finalement l'utilisation d'un module qui permette de rechercher et sélectionner des visualisations utilisable lors du processus d'analyse géovisuelle.

3.1.2.1. Formalisation des connaissances

Afin d'avoir un outil intelligent qui puisse guider les utilisateurs de l'information géographique dans l'exploration, l'analyse, et plus généralement l'utilisation de méthodes d'analyse géovisuelle, il est donc nécessaire d'avoir les éléments suivants :

- Une **formalisation de l'expertise du domaine d'intérêt** : les données du domaine, les motifs d'intérêt (menaces connues, comportements dangereux ou recherchés, etc.), les procédures habituelle pour l'analyse d'information, etc.
- Une **modélisation du profil d'utilisateur** : ses caractéristiques qui peuvent influer l'utilisation de la visualisation, la caractérisation de son environnement de travail (géographique, social et technique).
- Une **formalisation de l'expertise en géovisualisation / analyse géovisuelle** : les méthodes de visualisation d'information géographique, les interactions possibles, la forme visuelle des données, etc. En revenant sur le Reference Model Pattern (Heer et Agrawala 2006), cinq concepts doivent être pris en compte pour ce modèle.

Ces trois principales composantes permettent de développer un modèle de connaissances, en intégrant des cas d'étude et des méthodes de visualisation particulières. Dans notre cas, la formalisation des connaissances de ces différents domaines permet de mettre au point un métamodèle et des métadonnées sur l'utilisation de la visualisation d'information géographique, et l'étude des risques maritimes. Ces informations peuvent donc être étudiées pour comparer les besoins réels d'un opérateur, aux possibilités de l'analyse géovisuelle.

Ces modèles, ainsi que les éléments qui peuvent être déduits des modèles, forment donc une représentation « statique » de la connaissance grâce à la caractérisation des cas d'utilisation et des solutions d'intérêt pour un système socio-technique spécifique. Par la suite, nous allons donc appliquer ce modèle au domaine du *Maritime Domain Awareness*, qui comporte ses propres acteurs, outils et procédures que nous avons vu dans l'introduction de cette thèse. Dans la partie 3.1.3 de ce chapitre, nous présenterons notre méthodologie pour la formalisation de concepts qui modélisent les domaines identifiés.

3.1.2.2. Règles d'utilisation : possibilités et contraintes

Cette représentation statique de la connaissance doit alors être utilisée comme base pour le raisonnement et proposer des solutions à des situations données. Pour cela, des règles peuvent compléter ces modèles, en reliant ces trois composantes par des liens logiques. Celles-ci permettent en effet d'analyser les connaissances modélisées, interpréter des données en entrée et répondre à des questions et donner des conseils, selon la définition des systèmes experts données par Jackson (Jackson 1999).

Dans notre cas, le rôle des règles est donc de relier différents domaines de connaissances, en complétant l'expertise humaine relative à l'utilisation de la visualisation. Des règles d'utilisations possibles, ou *a contrario* de restriction, peuvent être développées selon les modèles. Ces possibilités ou restrictions d'utilisation de méthodes de visualisation peuvent, par exemple, se fonder sur le type de données et les règles de la sémiologie graphique. La Figure 3-3 illustre les trois composantes principales identifiées et le rôle des règles au sein d'un système à base de connaissances.

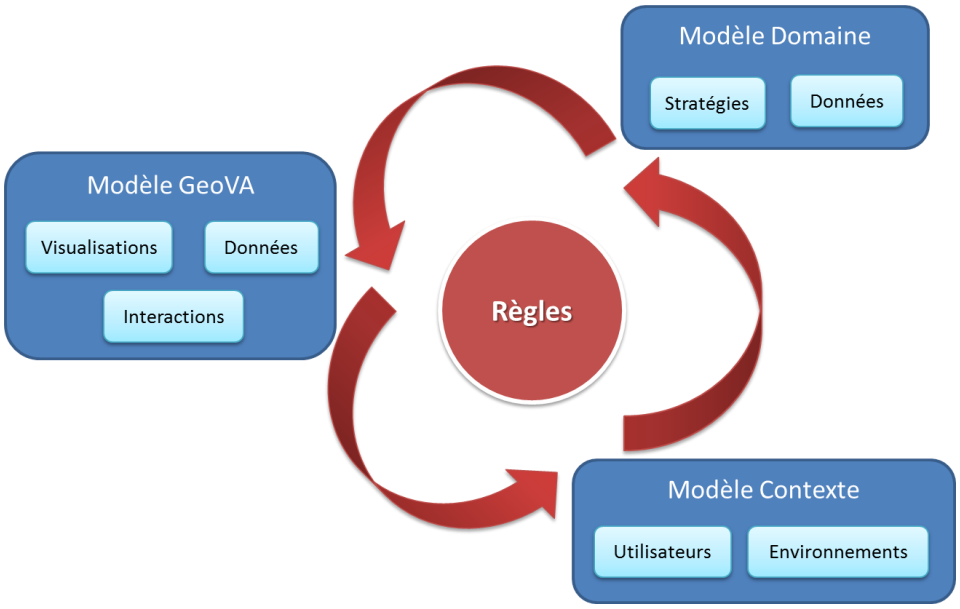


Figure 3-3. Modèle général pour l'étude des cas d'utilisation de l'analyse géovisuelle

Le formalisme des règles dépend la plupart du temps de la formalisation choisie et du langage correspondant, ou bien être basées sur un modèle très précis. Certains moteurs de règles s'appuient sur une modélisation par classes, d'autres moteurs et langage vont dépendre d'une modélisation ontologique RDF / OWL : c'est le cas des règles SWRL, telles que nous les avons présentées dans la partie 2.3.3 de cette thèse. Dans la partie suivante, nous expliquerons le développement et l'utilisation de règles, basées sur le modèle que nous proposons par la suite.

3.1.2.3. Module de sélection de visualisations adéquates

Tout comme les systèmes experts, les SAD et les PSE, un environnement d'aide à l'analyse géovisuelle s'appuie sur une modélisation des domaines d'intérêt, mais son utilisation se fait avant tout via une interface graphique, simple à utiliser et à comprendre. L'un des éléments principaux de notre EAAG est donc une interface de proposition de visualisations qui puissent être utilisées dans le cas d'utilisation étudié. Ce module doit présenter à l'utilisateur des méthodes qui sont proposées par le système à base de connaissances, tout en lui autorisant un choix dans les méthodes à utiliser.

Cette interface est notamment utilisée comme **module de sélection** de visualisations : c'est par cette interface que l'utilisateur peut interagir avec la base de connaissances, puis consulter les propositions qui lui sont faites.

Afin de développer un tel EAAG, de nombreuses possibilités techniques sont possibles. De manière à proposer un modèle qui soit à la fois partageable avec la communauté scientifique, qui puisse évoluer simplement et qui soit interprétable par des langages machines pour une utilisation intelligente de la connaissance, il a été nécessaire d'utiliser un formalisme respectant ces contraintes. De plus, il est nécessaire d'orienter la modélisation et le développement de ces concepts vers l'utilisation de règles, afin d'étudier la situation recontrée avec des contraintes et des lois du domaines de la géovisualisation. Le concept d'ontologie a été présenté dans les parties 2.3.1 et 2.3.2 de cette thèse, face à d'autres moyens de formalisations orientés objets et / ou règles. Par la suite, nous utiliserons donc ce mode de formalisation des connaissances, à travers des concepts, des relations et des axiomes.

La sous-partie suivante présente donc notre méthodologie de développement d'une ontologie pour la description des méthodes d'analyse géovisuelle, du domaine maritime et des cas d'utilisations pour le *Maritime Domain Awareness*.

3.1.3. Méthodologie de développement de l'ontologie

3.1.3.1. Choix de l'approche hybride

Afin de développer les ontologies qui composent l'EAAG, nous avons adopté une approche hybride, telle que décrite dans la partie 2.3.2.3 du chapitre 2. Cette approche permet d'effectuer une boucle de développement en partant des objets de haut niveau vers les détails de bas niveaux, puis remonter afin de contrôler la cohérence d'un tel modèle. Le domaine de la visualisation d'information de mouvement étant très large, nous avons procédé à cette approche hybride tout en limitant la diversité des concepts de plus haut niveau, et permettre son enrichissement par la suite.

Dans notre travail de recherche, l'approche descendante de la formalisation des connaissances a représenté un état de l'art et une étude des **environnements et des méthodes d'analyse géovisuelle** pour l'étude du mouvement et la caractérisation des objets mobiles et espace géographiques. Nous avons basé cette étude selon les modélisations proposées dans la littérature scientifique relative au domaine de la visualisation et des sciences de l'information géographique, que nous avons présenté précédemment (partie 2.2). Cette approche descendante permet de partir des concepts de haut niveau en visualisation, que sont les méthodes de visualisation, afin d'extraire les primitives qui servent à modéliser les visualisations étudiées. Ces primitives ontologiques, ou concepts, sont ensuite utilisées afin d'identifier les éléments principaux qui guident notre étude.

Cette première étape de formalisation nous permet donc de diviser les éléments de plus haut niveau, que sont les méthodes de visualisation, en éléments beaucoup précis et génériques que sont les concepts qui les constituent. Contrairement aux méthodes de visualisation, qui sont représentent des objets distincts les uns des autres, les concepts de plus bas niveaux qui les composent se retrouvent de nombreuses fois : les données, les abstractions visuelles, les tâches d'analyse, etc. L'approche descendante, ou *top-down*, permet donc un premier élagage des concepts et ainsi identifier les éléments les plus courants, à partir de visualisations sélectionnées.

Toutefois, avant d'identifier les concepts à développer, la première étape qui permet d'orienter le développement d'un tel modèle a consisté en l'identification des questions qui doivent être rendues possibles par l'utilisation de l'ontologie. Cette étape permet de limiter le spectre des concepts à

développer, ainsi que les règles à mettre au point. Dans cette sous-partie, nous présentons en premier lieu les contributions de l'ontologie au sein du système à base de connaissances pour un EAAG. Par la suite, nous présentons le procédé de développement des concepts, ainsi que les limites qui cadrent ce modèle.

3.1.3.2. Questions et utilisations d'une ontologie pour la formalisation de l'analyse géovisuelle

Dans un premier temps, lors du développement d'une ontologie, il est nécessaire de lister les premières requêtes qu'un tel formalisme doit rendre possibles. Ces requêtes peuvent être formulées dans un langage commun, qui permet alors d'identifier les objets fondamentaux d'une telle ontologie. Par la suite, des requêtes plus complexes peuvent alors être identifiées. Grâce à cette étape, des concepts, des relations et des attributs généraux peuvent alors être identifiés. Il est aussi nécessaire de fixer le périmètre d'utilisation d'une telle ontologie, en identifiant à la fois les problèmes qu'elle doit permettre de résoudre, et ceux qui ne font plus partie du domaine d'intérêt. Nous listons donc ici plusieurs requêtes qui doivent être résolues par l'EAAG, et qui doivent donc pouvoir être formalisées dans l'ontologie qui compose cet environnement d'aide à l'analyse.

Ces questions concernent plus généralement l'usabilité des méthodes d'analyse géovisuelle, afin de savoir lesquelles utiliser par qui, pour quoi et dans quelle situation. Pour cela, nous proposons un tableau qui décrit plusieurs types de questions :

- Les questions orientées maritimes, pour l'étude des mouvements de navires (domaine d'application de cette recherche)
- Les questions d'analyse géovisuelle plus générale, pour l'étude d'objets mobiles
- Les questions relatives à l'usabilité en général, soit le « *degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié* », concept défini par l'ISO 9241-11 (International Standards Organisation 1998)
- Les questions relatives à l'utilisation de la visualisation, en général
- L'usabilité relative aux différents profils d'utilisateurs et contextes

Dans la Table 3-1, nous présentons des requêtes possibles lors de l'utilisation d'une visualisation pour l'étude d'une zone maritime et des comportements des navires. A chacune de ces requêtes, les termes clés sont extraits, afin d'identifier les domaines que doit couvrir l'ontologie, pour la formalisation de ces questions et réponses.

Sans encore décrire les concepts et objets que l'ontologie doit intégrer, ces quelques exemples de questions mettent en évidence différents aspects à prendre en compte dans les connaissances que nous modélisons. Tout d'abord, si l'on suit le Reference Model Pattern, il est important de distinguer les données du domaine d'étude (dans notre cas, le domaine maritime), leurs propriétés intrinsèques (temporalité des données, dimensions, etc.), les données visualisables et les interactions que l'on peut avoir avec celles-ci.

Des scénarios d'intérêt pour le domaine doivent aussi être étudiés, tels qu'une dérive de navire, et le processus pour analyser et reconnaitre ce scénario doivent être pris en compte dans la base de connaissances. Cela permet d'intégrer l'approche humaine telle qu'elle est effectuée actuellement,

tel que Wills et Wilkinson ont pu le proposer grâce à leur logiciel AutoVis (Wills et Wilkinson 2010). Ainsi, des tâches d'analyse spécifiques au domaine étudié, ou bien plus générales dans l'étude des données de mouvement, doivent être intégrées dans la base de connaissance afin de caractériser les possibilités de la visualisation et les tâches analytiques à mener.

Table 3-1. Exemples de requêtes dans l'ontologie

Type de requête	Termes d'intérêt
Analyse maritime	
Quelle méthode permet de visualiser la position en temps réel des navires ?	Position (X, Y) Temps réel
Comment repérer une rencontre passée sur un espace défini ?	Rencontre Temps passé
Comment différencier un navire qui assiste, d'un navire qui attaque ?	Différencier
Quels navires vont passer dans une zone peu profonde ?	Profondeur Trajectoire
Quelles visualisations seront utiles pour analyser un comportement de pêche ?	Analyse de comportement
Analyse générale	
Quelles sont les zones à plus haute densité d'objets ?	Densité
Les données de mobilité présentent-elles un motif régulier ?	Motif
Quels sont les groupes d'objets à attributs proches ?	Regroupement Comparaison d'attributs
Utilisation générale	
Quelles méthodes permettent de représenter des données temporelles cycliques ?	Temps cyclique
Quelles données peuvent être représentées dans un cube spatio-temporel ?	Cube spatio-temporel
Quelles sont les interactions possibles pour une méthode choisie ?	Interactions
Usabilité	
Quelle méthode est la plus simple à utiliser pour un certain cas d'utilisation ? La plus complexe ?	Facilité d'utilisation
Quels utilisateurs peuvent facilement comprendre une carte auto-adaptative ?	Compréhension

Dans la prochaine partie, nous présentons la méthodologie de création des concepts et relations qui composent les ontologies des domaines étudiés. Pour cela, nous choisissons de mettre en œuvre la méthode hybride, que nous l'avons présentée dans la partie 2.3.2.3 de cette thèse, qui permet d'obtenir une ontologie aussi complète que possible.

3.2. Modélisation des concepts d'intérêt

3.2.1. Domaines couverts par l'ontologie

A partir de la définition donnée d'un EAAG dans la partie 3.1.1 de ce chapitre, nous identifions les différents domaines qui doivent être couverts par notre ontologie, afin de procéder à une approche descendante pour le développement des concepts. Rappelons les domaines qui doivent être pris en compte dans notre système à base de connaissances : d'un côté, le domaine de l'analyse géovisuelle, les méthodes et les lois qui régulent et orientent la visualisation de l'information ; d'un autre côté, le domaine des risques maritimes, les procédés d'analyse des mouvements pour leur identification et leur compréhension. Nous appelons ce second domaine le contexte d'utilisation, qui dépend du domaine d'application de l'analyse des informations géographiques (par exemple, le domaine maritime). Contrairement aux solutions d'analyse géovisuelle, dont la formalisation ne dépend pas du domaine d'application, la définition du contexte par les données, les problèmes et les utilisateurs dépend du domaine d'application. L'environnement d'aide à l'analyse géovisuelle comprend donc une ontologie générique, concernant les méthodes de visualisation, et une ontologie dépendant du domaine d'application.

Nous identifions ainsi plusieurs groupes généraux de concepts qui doivent être développés au sein de notre ontologie, pour l'étude des mouvements des navires en mer et leur comparaison à des possibles risques, menaces, etc. Suite à la définition donnée d'un EAAG, dans la partie 3.1.1, nous distinguons les domaines suivants :

- Analyse géovisuelle, méthodes de représentation de l'information de mouvement
- Acteurs susceptibles de jouer un rôle dans l'analyse géovisuelle des mouvements de navires
- Risques et scénarios possibles liés au trafic maritime
- Processus d'analyse et d'identification de ces risques
- Données maritimes liées aux mouvements des navires

Afin de pouvoir limiter le développement d'une telle ontologie, il est nécessaire de poser des limites pour la conceptualisation à mettre au point. Ces limites peuvent être identifiées au sein des futurs concepts, des règles et des requêtes possibles dans l'ontologie. Ainsi, nous nous limitons à :

- Une analyse des données de **mouvement** par des solutions **visuelles**, qui peuvent nécessiter plus ou moins de calcul machine (par définition de l'analyse géovisuelle)
- Des acteurs qui étudient les données de mouvement afin de comprendre et prendre des décisions : l'analyse des actions des contrôleurs et des capitaines de navires, par exemple, n'entre pas en compte dans le rôle de ce système à base de connaissances

La Figure 3-4 présente donc les trois grandes composantes de l'ontologie développé, afin de modéliser les processus et les solutions d'analyse géovisuelle pour l'étude du trafic maritime. Les parties suivantes présentent donc tour à tour ces différentes composantes, ainsi que les règles développées pour relier ces formalisations lors du raisonnement.

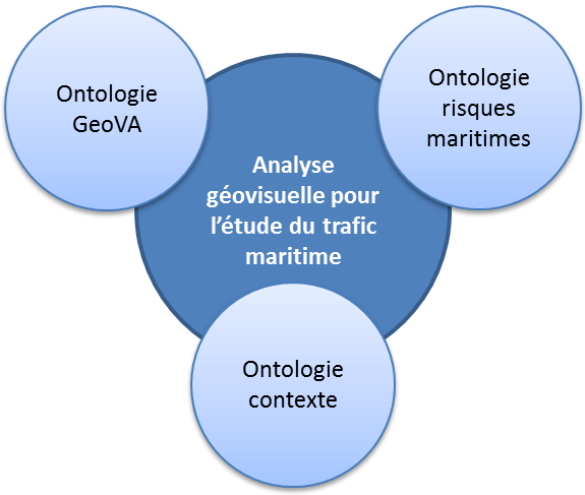


Figure 3-4. Domaines ontologiques couverts pour l'analyse géovisuelle du trafic maritime

3.2.2. Modélisation du contexte d'utilisation : le domaine maritime et la situation d'utilisation

La méthodologie générale du développement de l'ontologie ayant été présentée, nous présentons dans cette sous-partie les concepts et individus développés pour notre système à base de connaissances. Nous introduisons en premier lieu les concepts et individus relatifs au domaine maritime : ceux-ci doivent permettre de décrire le plus précisément possible les objets, les comportements et les situations d'intérêt pour l'analyse des mouvements de navire et l'identification des risques possibles.

3.2.2.1. Mouvements, données et risques maritimes

3.2.2.1.1. Données maritimes

A partir de l'étude de la structure des données AIS, les données utilisées en surveillance maritime ou en *Maritime Domain Awareness* le plus généralement, nous avons identifié plusieurs groupes de données, que nous regroupons sous le concept général *Data*. Notre étude se restreint à l'utilisation de l'analyse géovisuelle des données de mouvement : pour cela, nous avons limité les groupes de données conceptualisés à l'information relative au mouvement et son contexte. Les informations sur les opérations réalisées en termes de secours maritime, par exemple, ne concernent pas l'étude du mouvement et ne sont donc pas pris en compte dans ce modèle. Ainsi, nous retrouvons les éléments qui constituent les étapes successives de l'analyse géovisuelle du mouvement, selon Andrienko et Andrienko (2013) : les données intrinsèques aux objets, à leur mouvement et au contexte géographique où ils se déplacent. Pour cela, les données sont divisées en deux concepts de plus bas niveaux : ***VesselData*** pour les informations propres aux navires et leur mouvement, ainsi que ***BackgroundData*** pour les informations contextuelles.

La Figure 3-5 présente les *concepts* et les *instances* identifiés, ainsi que la hiérarchie qui les relie. Sur cette figure, extraite de l'éditeur Protégé 4.3, les symboles jaunes représentent les concepts, et les symboles violets représentent les instances. Il s'agit ici de la première étape de construction du noyau de l'ontologie, qui liste les concepts au sein de leur organisation hiérarchique. Une étape permettra, par la suite, de développer les relations sémantiques qui relient ces concepts aux autres

concepts de l'ontologie. Tel qu'elles sont représentées, les données relatives aux mouvements des navires peuvent être de plusieurs ordres :

- Les données dites « statiques » qui concernent l'identité des navires et les informations relatives au voyage effectué. Ces informations sont généralement transmises par les messages AIS. Malgré ce qualificatif, il est important de noter que ces informations peuvent être modifiées pendant le trajet du navire, comme la destination ou le statut de navigation.
- Les données de mouvement des navires, concernant leur cinématique. Celles-ci peuvent provenir des AIS émises, ou des capteurs de mouvement dont ils sont équipés, tels que le GPS, la centrale inertielle, etc.
- Les données qui ne sont pas diffusés par les modes de communication, mais dérivées des données collectées : les densités, les vitesses moyennes, l'ensemble des trajectoires, etc. Ces données représentent donc une première étape d'abstraction de l'information brute, telle que diffusée par messages AIS par exemple.

Pour les données contextuelles qui décrivent l'environnement de navigation, et qui peuvent donc être utilisées pour l'analyse du contexte et la prise de décision, on retrouve trois groupes majeurs :

- Les informations sur la règlementation maritime, telles que les couloirs de navigations, les zones de pêche interdite, les zones sensibles, etc. Ces informations sont notamment retrouvées dans les sources de données numériques telles que les ENC.
- Les informations sur le contexte maritime, qui ne concerne pas la régulation des zones de navigation, comme la bathymétrie.
- Les informations météorologiques qui décrivent l'état de la mer, du ciel, des courants, etc.

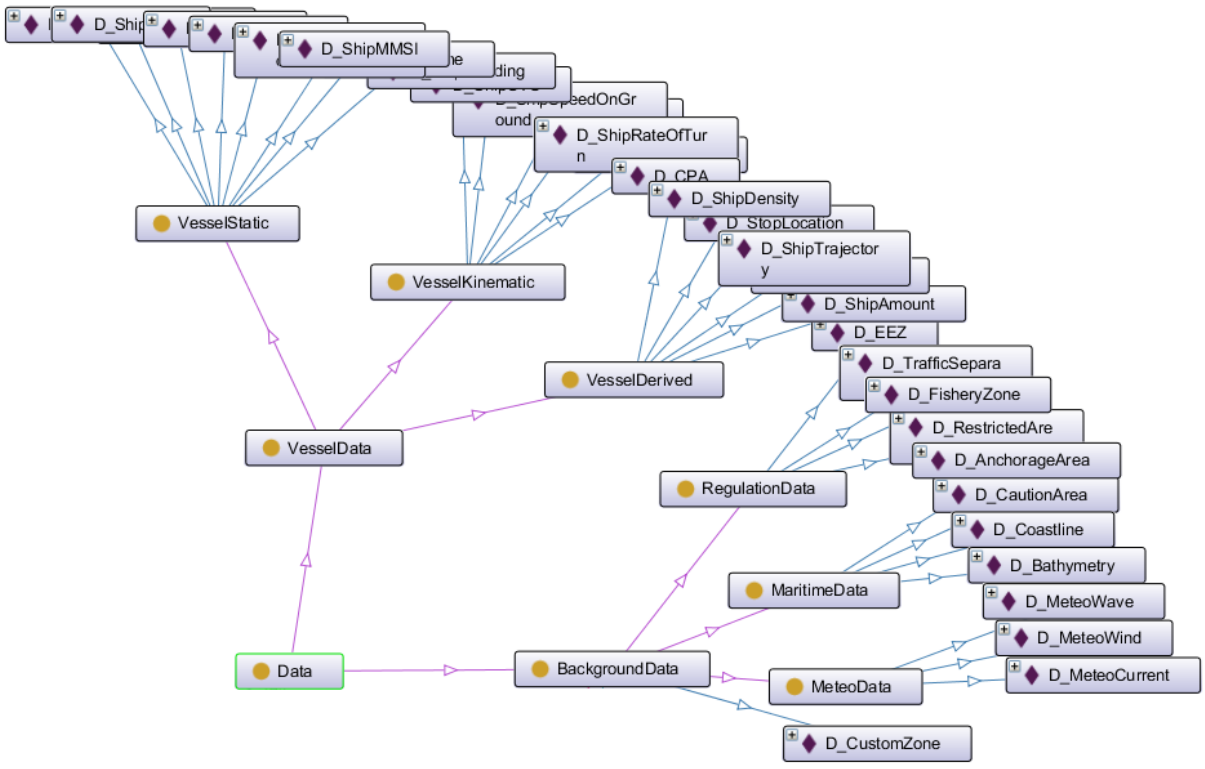


Figure 3-5. Hiérarchie de la typologie des données maritimes, dans l'éditeur Protégé 4.3

Chacun de ces concepts comporte alors de nombreuses instances de données concrètes, qui permettent de décrire l'espace maritime et les informations qui peuvent être projetées et visualisées pour une exploration et une analyse. Les relations sémantiques comportant ces concepts seront décrites à la fin de cette partie, une fois tous les éléments du cœur de l'ontologie présentés.

3.2.2.1.2. Temporalité des données

Comme nous venons de le décrire, certaines de ces informations peuvent être fixées pendant le déplacement d'un navire, ou bien une partie de cette route suivie (informations d'immatriculation IMO, informations relatives à la route, etc.) ; toutefois, d'autres informations peuvent être mises à jour de manière très régulière au cours des déplacements, afin de décrire les mouvements de l'objet. Comme nous avons aussi pu le voir dans l'étude de la manière de visualiser l'information, le temps est une dimension fondamentale à prendre en compte dans un environnement d'analyse géovisuelle. Il est donc nécessaire de savoir quelle est le format de la temporalité des informations étudiées, afin d'utiliser une méthode de visualisation adaptée. En effet, certaines méthodes de visualisation ne sont pas adaptées à des données qui évoluent en quasi-temps réel (e.g., des cartes de densité), et inversement.

Afin d'améliorer la description de cette hiérarchie de données, il est donc nécessaire de connaitre l'échelle temporelle de ces données, leur résolution temporelle, ainsi que leur *organisation temporelle*, à savoir si le motif représenté est cyclique (e.g., un trajet de bus) ou linéaire (e.g., les mouvements d'un piéton pendant 12h). L'aspect fondamental à prendre en compte dans notre cas, qui permet de différencier la surveillance (en temps réel) d'une analyse a posteriori (temps « historique »), est cette organisation temporelle de l'information. Pour cela, comme illustré sur la Figure 3-6, nous proposons un concept général de ***TimeDisplayed*** qui représente la temporalité des données étudiées, au sein de la visualisation. Ce concept peut alors être divisé en sous-concepts ***RealTime*** ou ***HistoricalTime***. Des informations passées enregistrées sur un laps de temps défini peuvent alors représenter une organisation cyclique ou linéaire, représentées par les deux derniers concepts de cette figure.

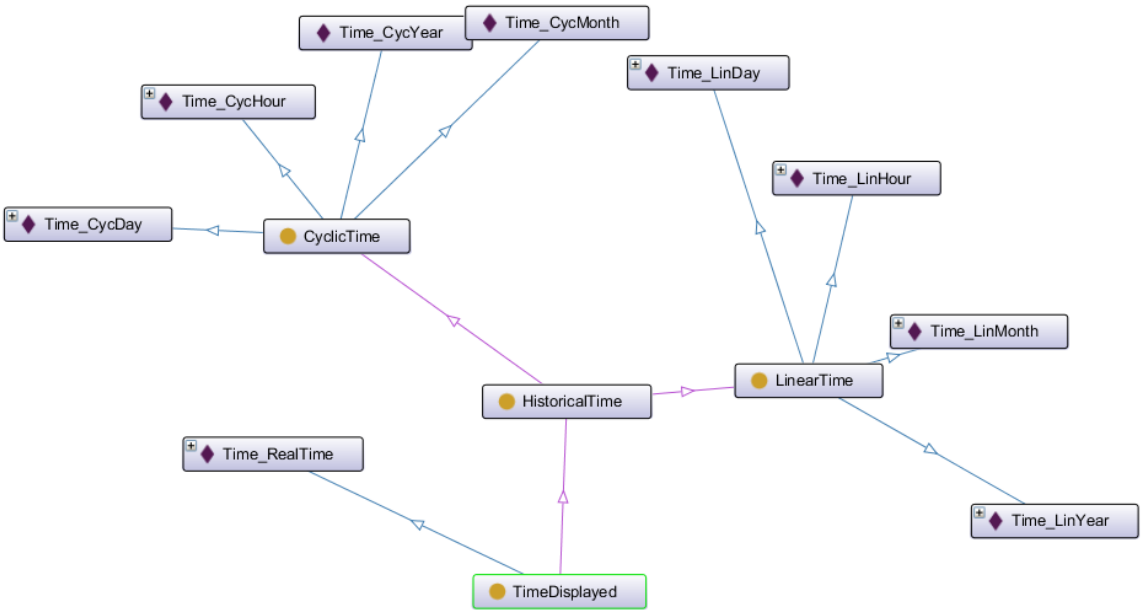


Figure 3-6. Caractérisation de la temporalité des données par les concepts et instances associées

Le développement de cette hiérarchie de concepts peut alors être approfondi vers des concepts temporels plus fins, en prenant notamment les échelles temporelles, les dates, etc. Toutefois, comme notre étude porte sur la comparaison de motifs qui peuvent avoir lieu à des dates différentes, mais utilisant pour cela une même méthode de visualisation, il n'est pas nécessaire que l'ontologie décrive un concept de temporalité trop complexe. Pour cela, nous choisissons le niveau le plus fin de cette hiérarchie (i.e., les instances de l'ontologie) comme étant description de la résolution temporelle des données.

Une information en temps (ou quasi-) temps réel ne peut pas être subdivisée en instances similaire, la notion de temps réel induisant une résolution temporelle assez fine. La temporalité des données, concernant essentiellement le déplacement des navires, n'a donc aucun intérêt à être qualifiée d'annuelle, mensuelle, ou encore quotidienne : les données de mouvement des navires sont en effet émises de manière régulière, à une résolution proche de la minute. Diviser le concept *RealTime* en plusieurs instances perd donc son intérêt dans la caractérisation des données. En terme d'ontologie, comme le dernier objet d'une hiérarchie doit être une instance, il est donc plus intéressant de considérer l'objet *RealTime* comme une instance, et non pas un concept de l'ontologie.

Les concepts de données et de temporalité décrivent le corpus des informations qui peuvent être utilisées dans l'exploration et l'analyse des données de mouvement, dans le but d'identifier et comprendre les risques maritimes. Ceux-ci peuvent aussi être décrits de manière très précise, tel que nous le présentons dans la sous-partie suivante.

3.2.2.1.3. Risques maritimes

Les risques maritimes, qui sont la raison de l'analyse du trafic maritime, peuvent être formalisés de nombreuses manières. De nombreux travaux ont porté sur la description formelle des anomalies et des comportements des navires, pour la reconnaissance automatique des risques possibles et la levé d'alertes. Nous pouvons citer les travaux de Roy (Roy 2008 ; Roy 2009 ; Roy et Davenport 2009) et Vandecasteele (Vandecasteele et Napoli 2013 ; Vandecasteele et al. 2014) pour la formalisation des

anomalies maritimes : ceux-ci se basent notamment sur la formalisation des mouvements et des relations topologiques entre les objets mobiles.

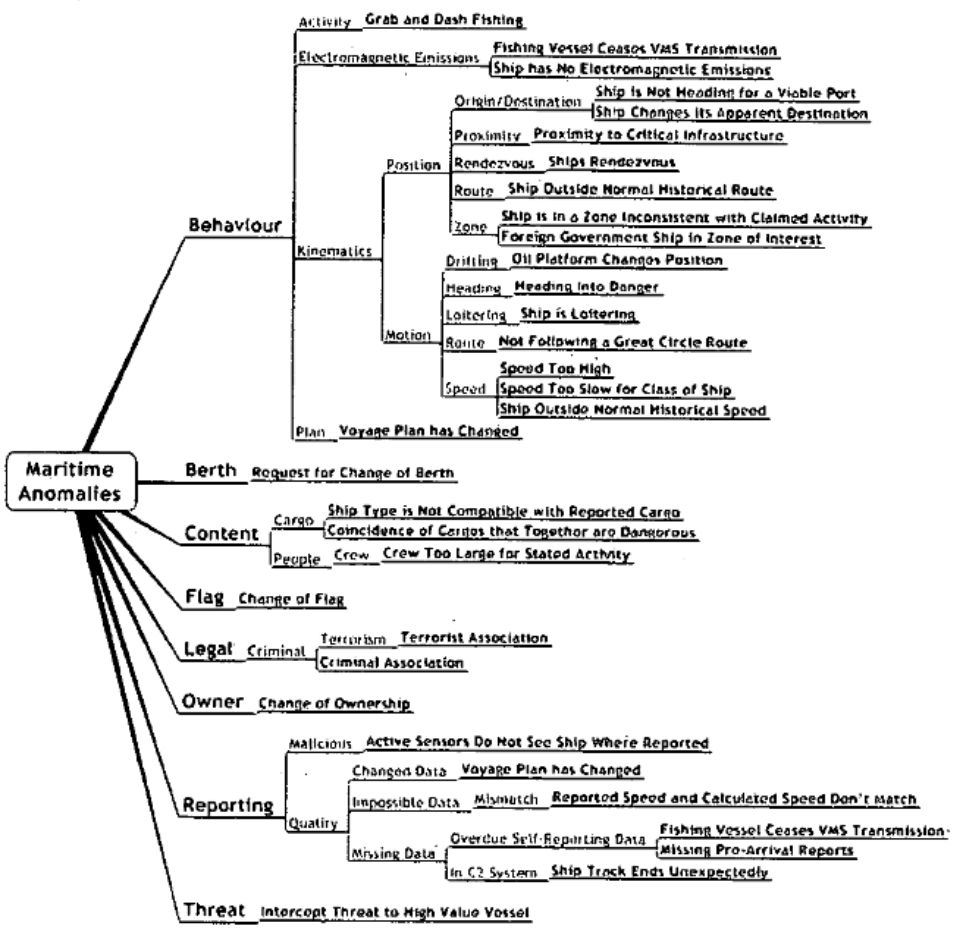


Figure 3-7. Catégorisation des anomalies maritimes, selon la taxonomie des faits (*source :* Roy et Danveport 2009)

Par de nombreux entretiens avec des experts du domaine maritime, Roy étudie des anomalies telles que la pêche illégale, l'immigration illégale, l'attaque de pirates, l'entrée dans des zones dangereuses ou régulées, etc. Afin de les caractériser, une taxonomie de **faits maritimes** est développée pour décrire les possibles comportements et relations entre des objets X, Y et leur environnement. De ces faits maritimes, une taxonomie des **anomalies maritimes** est alors extraite et décrite par des **règles de détection** (Roy et Davenport 2009). Cette taxonomie des risques et anomalies (voir la Figure 3-7) comporte à la fois des anomalies sur la dynamique et le comportement du navire (vitesse, localisation, trajectoire, manœuvre, etc.) et des anomalies non-cinématiques (état de l'AIS, liste des passagers, contact avec les ports, etc.).

Dans ses travaux, Vandecasteele (2012) propose quant à lui une approche ontologique pour la formalisation des comportements, par une caractérisation de la dynamique, de la forme et de la topologie des mouvements. Des règles sont mises en place, basées sur la description de ces comportements, de manière à détecter automatiquement des comportements similaires à des scénarios de risques et d'anomalies définis.

Ces travaux montrent que la modélisation de risques n'est pas considérée comme un ensemble de concepts et relations, mais un ensemble de règles qui s'appuient sur la formalisation des comportements. D'une manière similaire, notre recherche propose d'anticiper les risques non pas par une modélisation formelle de ceux-ci, mais par une succession de tâches d'analyse, d'exploration. Ceux-ci représentent la logique de travail des analystes ou des contrôleurs dans la compréhension des données de trafic, pour la détection de risques et d'anomalies.

A titre d'exemple, nous avons donc utilisé un ensemble de risques issus des travaux de Roy et de Vandecasteele, sans formaliser les comportements qui leurs sont associés. De plus, par le sujet de notre étude, seuls les risques et anomalies relatifs à la dynamique des navires sont pris en compte. Au sein de notre ontologie, il s'agit donc d'un unique concept ***RiskType***, qui comporte plusieurs instances de risques possibles, tel que l'illustre la Figure 3-8.

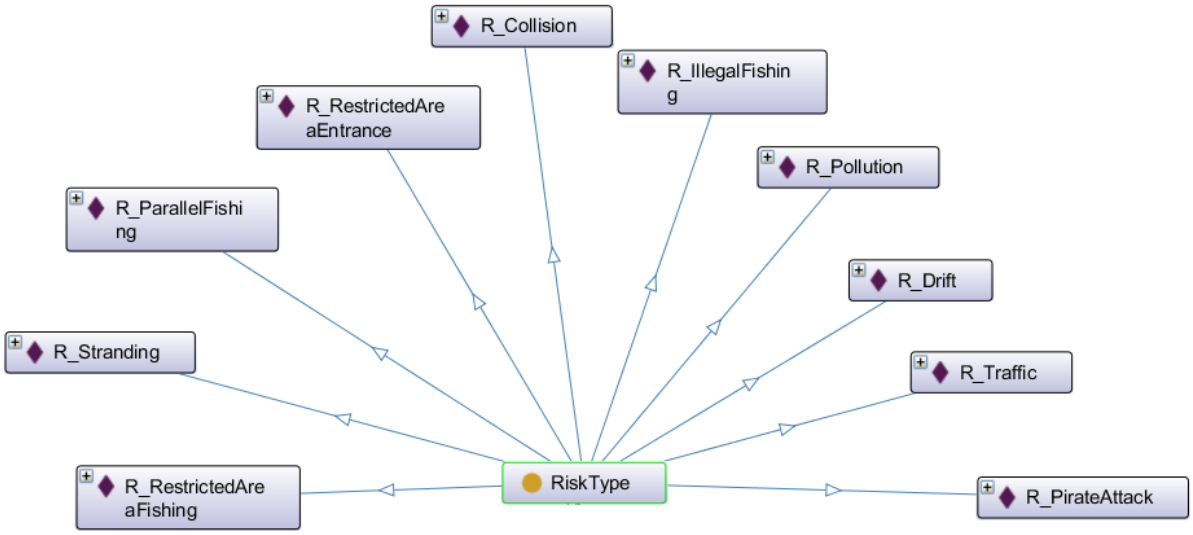


Figure 3-8. Concept de risques et instances proposées dans l'ontologie

Nous présentons maintenant notre travail de hiérarchisation pour les concepts qui décrivent les tâches d'analyse, afin de modéliser le processus d'analyse des risques qui peut être mené par l'utilisateur.

3.2.2.2. Tâches d'exploration et analyse des comportements

Nous avons vu dans la partie 1.3.2 les différents procédés d'analyse du mouvement, grâce à la visualisation d'information. Les grandes étapes ont ainsi été présentées, ainsi que les analyses plus spécifiques qui les composent. Dans notre cas, l'exploration et l'analyse des données de mouvement concernent donc les trajectoires de navires au sein du contexte géographique, règlementaire et météorologique de la mer. Au cours de l'étude de ces trajectoires, de nombreuses questions (tâches) sont posées, par les contrôleurs maritimes ou les analystes. Ces questions d'exploration peuvent être ordonnées des plus générales (identification) aux plus précises (interactions entre objets et contexte), tel que ce proposent les méthodes présentées par Andrienko et Andrienko (Andrienko et Andrienko 2013). Ces étapes générales sont ainsi reprises sous forme de concepts, tel qu'illustré sur la Figure 3-9. Ces groupes correspondent aux trois étapes successives dans l'analyse géovisuelles de données spatio-temporelles, auxquelles nous avons ajouté l'étape finale d'identification des risques. Ces concepts sont regroupés sous le concept général de ***Task***, qui correspond donc à une tâche

d'analyse géovisuelle. Nous avons présenté cette classification ontologique utilisée dans notre base de connaissances et son utilisation au sein de règle générales, lors de conférences au cours de notre travail de recherche (Vatin et Napoli 2013d ; Vatin et Napoli 2013e).

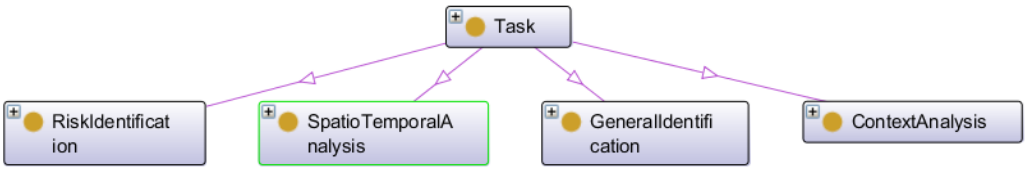


Figure 3-9. Typologie des tâches d'exploration et d'analyse

Rappelons donc la signification de ces quatre concepts, appliqués à l'étude des données maritimes. L'**identification générale (*GeneralIdentification*)** est la première étape dans l'exploration des données. Elle consiste à retrouver des objets d'intérêt depuis des motifs connus (par exemple, extraits de data mining) ou identifier des objets (reconnaitre le nom et les attributs des objets sélectionnés). L'**analyse spatio-temporelle (*SpatioTemporalAnalysis*)** permet d'analyser la cinématique, les arrêts et la trajectoire des objets mobiles. Les évènements passés et les comportements des objets vont être expliqués en comparant les objets et leur contexte, dans le temps et l'espace. L'**analyse du contexte (*ContextAnalysis*)** regroupe la recherche d'informations qui puissent caractériser le contexte géographique et son interaction avec les données étudiées. Ce contexte peut être des informations météorologiques, bathymétriques, ou bien sur la dangerosité de la zone (attaques passées de pirates, densité de navires, etc.). L'**interprétation des risques (*RiskIdentification*)** finalise l'étude des objets mobiles. On recherche une possible corrélation entre l'analyse spatio-temporelle est l'analyse du contexte, afin d'expliquer un évènement passé, ou ce qu'il peut se passer.

De très nombreuses taxonomies des tâches analytiques, par la visualisation d'information géographique, ont pu être mise au point cette dernière décennie. Roth (Roth 2012) a notamment pu réaliser un état de l'art de ces différentes tâches, qu'il appelle les primitives de contrôle « orientées objectif » (*objective-based*). Son étude permet d'identifier, au sein d'un large corpus d'articles, les termes génériques qui sont le plus souvent utilisés. La cartographie de ces résultats est illustrée sur la Figure 3-10, par des liens de proximité entre termes, ainsi que leur importance dans les textes étudiés. Cette représentation des primitives de contrôle correspond donc à une étape de formalisation des tâches d'analyse que doit mener l'utilisateur.

Afin d'aboutir à cette cartographie des primitives de contrôle, Roth souligne l'importance d'étudier les séquences d'interactions passées lors de l'utilisation de la géovisualisation : bien que le but général d'utilisation de la carte soit généralement connu, la manière d'y accéder et les tâches successives ne sont pas nécessairement claires pour l'utilisateur. L'étude de Balzarini et al. (Balzarini et al. 2013) montre la grande diversité des processus d'analyse et d'exploration mis au point, selon des experts de l'IG ou des non-experts.

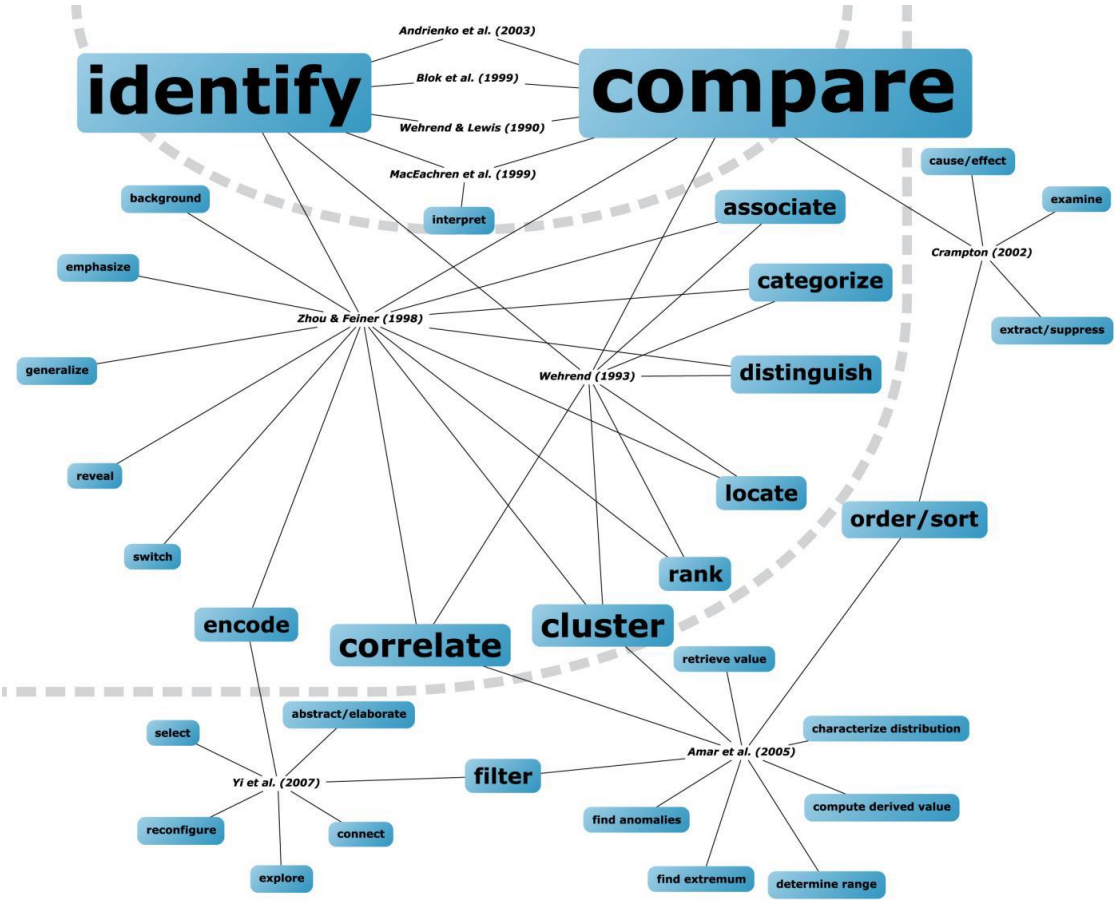


Figure 3-10. Cartographie des primitives de contrôle *orientées objectif* (*source :* Roth 2012)

Au sein de cette cartographie générale, Roth identifie trois sections d'objectifs : le premier groupe contient les deux tâches fondamentales d'identification et de comparaison, qui sont très régulièrement retrouvées dans la littérature. Il s'agit des deux tâches d'exploration de base, qui forment notre première manipulation lors de l'étude de données. Ces termes sont de plus haut niveau : d'autres tâches d'exploration et d'analyse plus spécifiques peuvent alors être ajoutées, de manière plus détaillée.

Ces actions plus spécialisées sont représentées dans la partie centrale du schéma, grâce à des termes plus techniques qui peuvent faire appel à des méthodes spécifiques de visualisation et d'analyse de l'information. Ainsi, on retrouve des techniques telles que le regroupement d'objets (clustering), la corrélation, la classification, etc., qui permettent de mener les tâches globales d'identification et de comparaison.

Enfin, la partie inférieure du schéma présente des termes moins partagés par la communauté scientifique. Roth interprète cet effet d'isolation comme venant du fait que ces termes sont plus explicites en tant qu'opérateur (manipulation des données), que des objectifs d'utilisation proprement parler. Nous aurons l'occasion de revenir sur les opérateurs dans la partie 3.2.3.3 de ce chapitre.

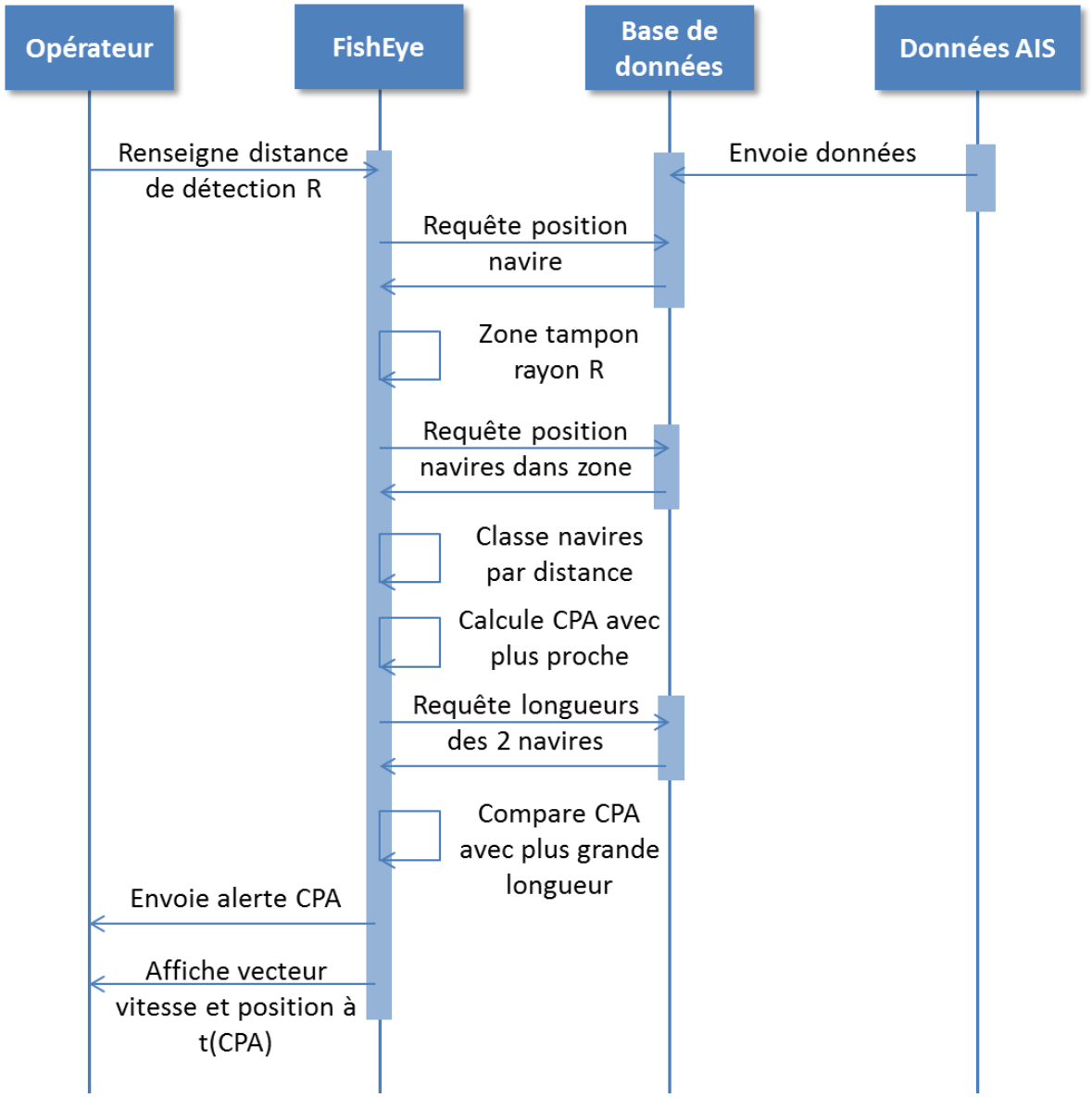


Figure 3-11. Utilisation du concept de Closest Point of Approach pour la prédiction d'une collision entre deux navires (Cario et Gravière 2013)

Afin d'appliquer cette étude au domaine maritime et de comparer les concepts identifiés par Roth, nous pouvons effectuer une étude des questions qui se posent lors de l'analyse des déplacements d'un navire, afin de rechercher une anomalie, ou un possible risque. Pour cela, la Table 3-2 présente un exemple de scénario d'analyse, ainsi que les primitives pour les tâches correspondantes. Cet exemple a pu être été approfondi grâce aux travaux d'élèves officiers de la Marine Nationale, notamment pour les scénarios de collisions, d'échouement et de pêche illégale (Cario et Gravière 2013). La Figure 3-11 introduit l'utilisation du point de plus proche approche (CPA) pour l'étude d'une possible collision entre deux navires, au sein de données de mouvement, pour illustrer ce processus d'analyse par la visualisation, assisté par des calculs automatisés. Les primitives présentées dans le tableau ne sont donc pas des *interactions* proposées par la visualisation, mais des *questions* (tâches) posées lors de l'utilisation. Certaines ont pu être dérivées des primitives de Roth, de façon à être intégrées aux concepts du processus d'analyse visuelle qui ont été identifiés précédemment. Ces

questions d'analyse sont donc organisées selon ces concepts, afin de suivre le processus d'analyse proposé par Andrienko et Andrienko (2013).

Table 3-2. Exemple de questions d'exploration et d'analyse visuelle pour la prédiction de collision entre deux navires

Questions d'analyse	Primitive correspondante
Identification générale	
Localiser un navire à analyser	Localiser un objet
Identifier un navire à analyser (type, destination, etc.)	Identifier un objet
Analyse du contexte	
Les navires se voient-ils ?	Caractériser l'environnement
Y a-t-il des navires proches ?	Caractériser l'espace
Est-ce un axe important de navigation ?	Regrouper, Synthétiser
Analyse spatio-temporelle	
Quels navires ont une vitesse non nulle ?	Comparer les dynamiques, Mesurer la vitesse
Les trajectoires vont-elles se croiser ?	Prédire une trajectoire
A quelle distances les navires vont être le plus proche ?	Mesurer l'espace

Les figures suivantes (Figure 3-12, Figure 3-13, Figure 3-14, Figure 3-15) reprennent alors les diverses primitives d'utilisation des cartes, qui ont pu être identifiées pour les deux cas présentés. Les quatre groupes de primitives sont séparés, afin de mettre en évidence le processus d'exploration de l'information géographique. Chacune de ces primitives de tâches modélisées est décrite avec plus e précisions dans les annexes de ce manuscrit (voir l'annexe *Description des tâches modélisées*, p.231). Lorsque ces tâches devront être effectuées, de nombreuses caractéristiques décrivant l'utilisateur et son environnement de travail vont directement influencer sur l'utilisation des environnements d'analyse géovisuelle. La prochaine partie concerne donc la caractérisation du **contexte d'utilisation**.

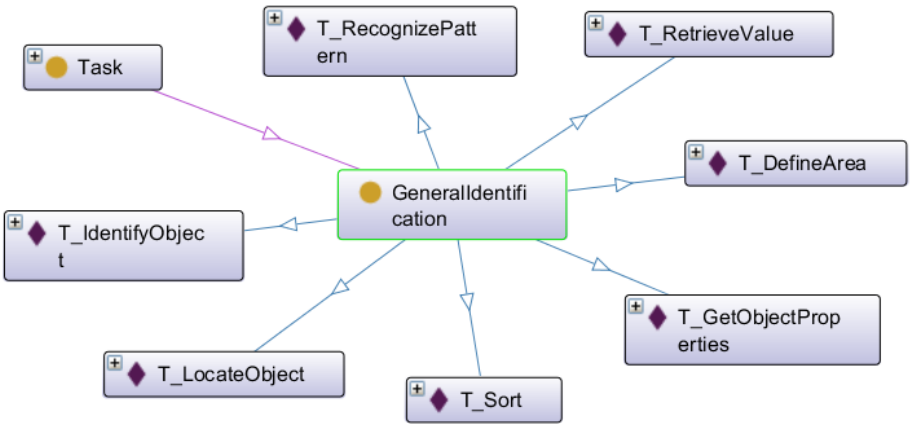


Figure 3-12. Tâches d'identification générale pour l'étude du trafic maritime

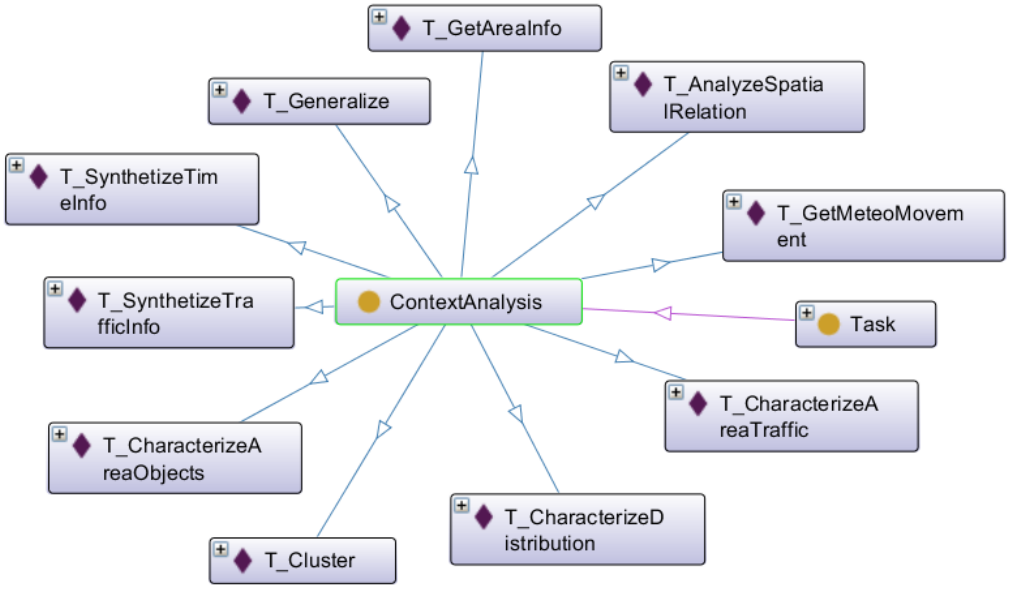


Figure 3-13. Tâche d'analyse du contexte pour l'étude du trafic maritime

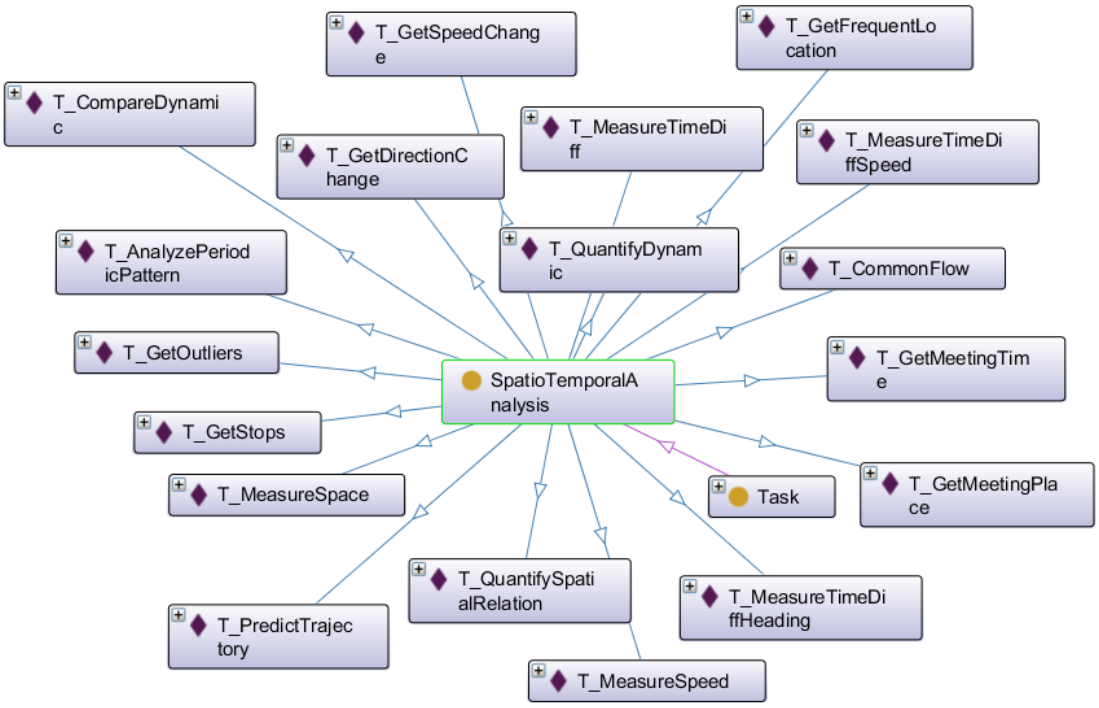


Figure 3-14. Tâches d'analyse spatio-temporelle pour l'étude du trafic maritime

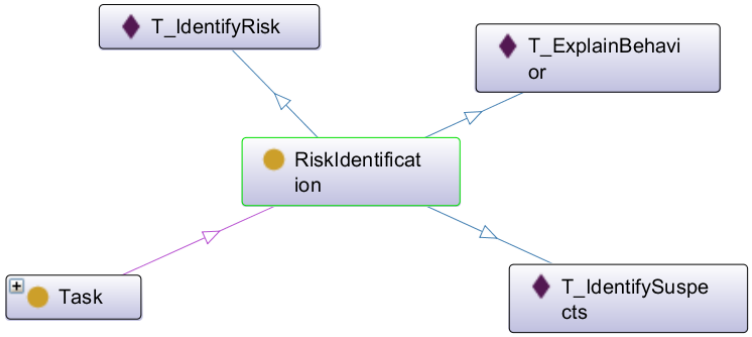
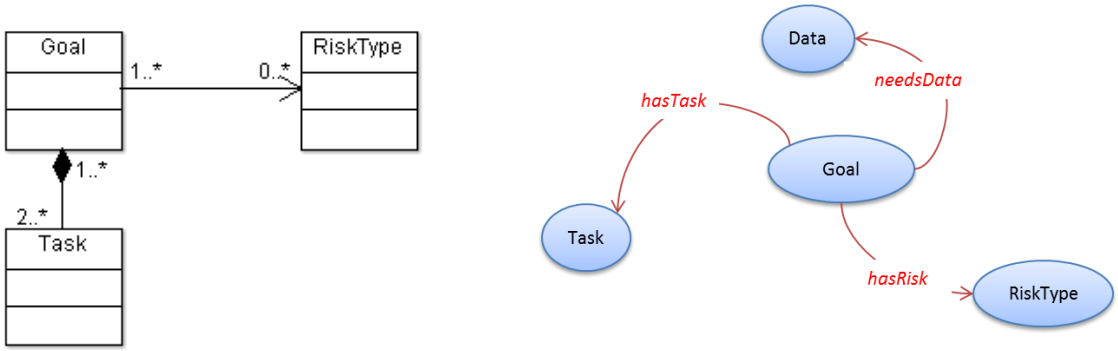


Figure 3-15. Tâches d'identification du risque pour l'étude du trafic maritime

Grâce à ces termes, nous pouvons caractériser une **stratégie de visualisation**. Nous définissons cette stratégie comme un enchaînement de tâches, dans un but d'exploration de la donnée (recherche de motifs) ou d'analyse (validation d'une hypothèse). Dans le cas de l'analyse, l'hypothèse émise peut être relative à un risque spécifique ou non, que l'utilisateur souhaiterait confirmer, par exemple. Ainsi, nous introduisons le concept de *Goal*, ou but d'utilisation.

Le concept *Goal* peut dépendre d'un vocabulaire spécifique au domaine d'étude et correspondre à une liste ordonnée de tâches, ou bien être une unique tâche à opérer sans lien à un risque. Ainsi, ce concept peut être défini tel qu'illustré sur la Figure 3-16 (a) par le formalisme UML, ou sur la Figure 3-16 (b) correspondant au concept dans l'ontologie.

La succession de ces tâches d'analyse, selon le type de risque recherché, a pu être étudié au cours de travaux d'élèves officiers de la Marine Nationale encadrés par notre équipe (Cario et Gravière 2013). Dans ces travaux, le processus d'analyse des trajectoires, ainsi que les notions maritimes nécessaires à la compréhension du trafic, ont pu être modélisés par le langage UML. Certains scénarios ont ainsi pu être étudiés en précision, selon les tâches analytiques qui leur correspondent ainsi que les possibles calculs à effectuer (par exemple le point de plus proche approche, la sonde minimum de passage, etc.).



(a) (b)

Figure 3-16. Le concept Goal, représenté par UML (a) et par une version sémantique (b)

Le domaine d'application et les données correspondantes ayant été présentés par les concepts d'intérêt, nous présentons maintenant les différents profils d'utilisateurs qui doivent intervenir dans l'étude des données de trafic maritime, ainsi que la manière de caractériser leur profil.

3.2.2.3. Utilisateurs de la visualisation de données de mouvements

Au cours de nos recherches sur le domaine maritime et la visualisation d'information, nous avons identifié plusieurs personnes susceptibles de travailler sur des données de trafic maritime, en utilisant comme base de travail la visualisation d'information (plus ou moins avancée) et les cartes. Parmi ces nombreux acteurs de la sécurité et sûreté maritime, nous avons identifié les rôles principaux lors la présentation du contexte maritime en introduction de ce manuscrit, ainsi que par les travaux de Wybo (Wybo 2012) présentés dans la partie 1.1.1.3. Ainsi, nous pouvons citer plusieurs exemples de profils distincts :

- les personnes à bord des navires, tels que les **capitaines** et **officiers** de passerelle, qui sont au cœur même des mouvements des navires. Leur interprétation des données diffère d'un point de vue « sur le continent », car ils n'ont pas la même perception des mouvements et des dangers qu'un observateur externe.
- les **contrôleurs**, dans les centres de surveillance tels que les CROSS, que nous les avons présentés précédemment. Ces opérateurs doivent surveiller le trafic en temps réel depuis la côte, afin de repérer des anomalies, guider les navires pour la résolution de leurs problèmes, ainsi que coordonner les actions d'intervention et de sauvetage en mer.
- les **analystes et experts**, dans les bureaux d'études tels que le BEAmer en France (bureau d'enquêtes sur les évènements de mer), le MAIB en Angleterre (*Marine Accident Investigation Branch*), etc. Ces experts ont pour buts de rechercher et déterminer les causes techniques des accidents en mer. Suite à ces missions, les rapports d'analyses sont publiés en ligne pour une consultation publique.
- les **chercheurs** du monde académique, qui sont amenés étudier les données maritimes tels que les mouvements et les accidents passés. Leur rôle majeur est de modéliser l'information étudiées et proposer de nouveaux modèles de risques, afin d'améliorer le contrôle du trafic maritime et la gestion des dangers en mer.

Ces nombreux utilisateurs de l'information géographique pour la gestion des risques en mer sont alors amenés à utiliser des méthodes de visualisation très différentes, selon leurs habitudes, leur expérience et les outils de leur environnement de travail. Afin de les différencier, nous proposons plusieurs critères d'intérêt pour cette étude, qui peuvent alors être formalisés dans un « profil d'utilisateur » au sein de l'ontologie du domaine d'étude.

3.2.2.3.1. Caractéristiques de l'utilisateur

Les utilisateurs qui sont amenés à étudier l'information géographique, dans la gestion des risques maritimes, peuvent avoir différentes professions, donc différents niveaux et domaines d'études (notamment en statistiques et informatique, domaines clés pour l'analyse géovisuelle). De plus, ces différents acteurs peuvent avoir différents niveaux d'expérience dans le domaine du trafic maritime et des risques qui lui sont associés. Nous pouvons d'ailleurs remarquer que la « connaissance du domaine » (données, évènements, risques, etc.) est l'un des trois axes du cube de la (Cartographie)[3] de MacEachren (1995), qui définit l'utilisation de la visualisation d'information géographique. Cette notion de connaissance des données se retrouve aussi lors de l'étape *Vigilance*, définie par Wybo dans le processus de gestion des risques (Wybo 2012) : certains risques sont connus et déjà définis (phase de modélisation), auquel cas l'utilisateur sait quoi rechercher dans les visualisations

d'information ; ou bien peuvent être nouveaux et nécessiter une phase d'analyse et de modélisation importante.

Au travers de ces études précédentes, nous pouvons donc identifier trois premières caractéristiques d'intérêt pour cette étude : le ***rôle*** de l'utilisateur (ou sa profession), son ***expérience*** dans le domaine d'application (connaissance des scénarios, des comportements habituels, etc.) et son ***éducation scientifique***, c'est-à-dire son habitude d'utilisation d'outils scientifiques tels que les statistiques, ou une abstraction avancée des informations visualisées. L'utilisation faite de la visualisation dépend aussi fortement du contexte dans lequel évolue et travaille l'utilisateur (Mallé-Noyon 2008 ; Lopez-Velasco et al. 2009 ; Luqman et Griss 2010), comme nous l'avons décrit lors de la définition d'un EAAG : cela peut provenir de la mobilité ou non de son environnement de travail (système embarqué sur un navire), ou bien encore des conditions technologiques, sociales et culturelles dans lequel travaille l'opérateur. Pour cela, nous avons porté notre choix de formalisation sur deux aspects à traiter : les ***compétences technologiques*** de l'utilisateur, qui dépendent de son confort/habitude avec la visualisation d'information, les possibles formations suivies, mais surtout son cadre de travail. En effet, si les machines utilisées pour une analyse d'informations de mouvement ne sont pas assez puissantes, les méthodes nécessitant de grandes capacités de calcul et d'affichage graphique seront impossibles à réaliser.

Afin d'estimer les caractéristiques qualitatives liées à la description de l'utilisateur, nous choisissons de limiter les valeurs possibles : celles-ci sont des informations qualitatives ordonnées. Pour cela, nous utilisons une échelle à trois niveaux (*Low* < *Medium* < *Advanced*) pour chacune des caractéristiques ordinales. Cela permet ainsi d'implémenter des utilisateurs-type dans l'ontologie, par la suite. La Table 3-3 reprend les caractéristiques que nous avons identifiées. La sous-partie suivante présente comment les modéliser au sein de l'ontologie.

Dans notre étude, nous définissons donc un potentiel utilisateur de visualisation comme une personne (de sexe et âge indistincts), dont les capacités en termes d'analyse des données (du domaine maritime, ici) sont caractérisées tel que nous l'avons présenté dans la sous-partie précédente. Afin de développer des concepts dans l'ontologie, cela ne correspond pas à une hiérarchie de niveaux, mais à des liens sémantiques qui relient l'utilisateur à ces caractéristiques.

Table 3-3. Description des caractéristiques retenues pour la description de l'utilisateur

Caractéristique	Description	Exemples
Rôle	Profession de l'utilisateur, permet d'identifier les tâches potentielles et les valeurs de certaines caractéristiques.	Contrôleur, Analyste, Chercheur
Education scientifique	Permet d'évaluer les compétences en statistiques, mathématiques.	Faible, Moyen, Avancé
Expérience du domaine	Donne le niveau de connaissances avec le domaine maritime : données, scénarios possibles, etc. Il s'agit de qualifier son expérience.	Faible, Moyen, Avancé
Compétences technologiques	Permet d'évaluer les compétences en informatique et visualisation.	Faible, Moyen, Avancé
Environnement de travail	Décrit le milieu où évolue et travail l'utilisateur, si celui-ci est mobile, etc.	Centre de surveillance, Centre de recherche, Navire

3.2.2.3.2. Rôle du temps dans cette étude

Dans la première partie de ce chapitre sur la modélisation, nous avons introduit le concept relatif à la temporalité des données. Il est toutefois important de noter que la temporalité des données peut être distincte du « temps de traitement » de ces mêmes données: l'étude de données en temps réel, s'il n'y a pas de danger ou de risque, peut être effectuée sur un temps plus long que l'étude de données en temps réel qui sont relatives à un danger avéré. De façon inverse, l'analyse de données passées n'est pas nécessairement synonyme d'un temps de traitement long : certaines contraintes peuvent nécessiter des résultats d'analyse rapides. Cette question fait donc appel à la notion de pression, qui peut être due à un temps court pour le traitement, l'analyse des données : la pression exercée par cette étape peut alors avoir un impact majeur sur la façon dont les cartes et les visualisations sont utilisées et comprises (Wilkening et Fabrikant 2011).

Toutefois, cette pression opérée par un temps relativement court pour l'analyse des données est difficilement quantifiable et modélisable, par rapport à notre problématique de recherche. De plus, la pression que subit l'utilisateur ne peut pas être considérée comme une caractéristique propre à la personne (profil utilisateur), mais plutôt comme dépendante de son usage dans une situation spécifique. Par la suite, nous ne prenons pas en compte ce facteur externe, et nous considérons la contrainte de temps de traitement comme nulle. Ainsi, nous conservons la seule temporalité des données afin de caractériser le cas d'utilisation.

La sous-partie suivante présente le développement de tels concepts et relations afin de représenter un utilisateur de la visualisation d'informations géographiques et maritimes, au sein de l'ontologie.

3.2.2.3.3. Concepts ontologiques pour le modèle utilisateur

Les valeurs des trois caractéristiques *EducationLevel*, *ExperienceLevel* et *TechnologicalAbilities* sont alors notées comme un entier entre 1 et 3, 1 correspondant au niveau Faible. Les rôles et environnements de travail peuvent, quant à eux, être décrits par des concepts, qui regroupent eux-mêmes un ensemble d'instances possibles. La Figure 3-17 représente un extrait de la modélisation de l'utilisateur au sein de l'ontologie. Ce modèle a été présenté en workshop suite à des travaux communs à l'ITC (University of Twente) (Vatin et Napoli 2013b). Les instances possibles de *WorkEnvironment* et de *Role* ne sont pas entièrement développées : l'ensemble des possibilités est trop vaste et peut facilement s'éloigner de notre question de recherche d'origine, qui concerne la modélisation des environnements d'analyse géovisuelle.

Sur cette figure, les instances de rôles telles que *Controller* ou *Analyst* représentent des rôles prédéfinis, qui pourront servir d'exemple par la suite. Ces rôles sont eux-mêmes qualifiés par l'environnement de travail et les caractéristiques de l'utilisateur. En distinguant les concepts de *Role* et de *User*, notre idée est de permettre à un même utilisateur de changer de rôle : si son expérience dans le domaine maritime augmente, ou bien s'il suit une formation pour un logiciel ou une visualisation, ses caractéristiques changeront, et peut-être son rôle dans la boucle décisionnelle. La Table 3-4 présente des profils types d'utilisateur qui ont été intégrés dans l'ontologie, en suivant le formalisme présenté.

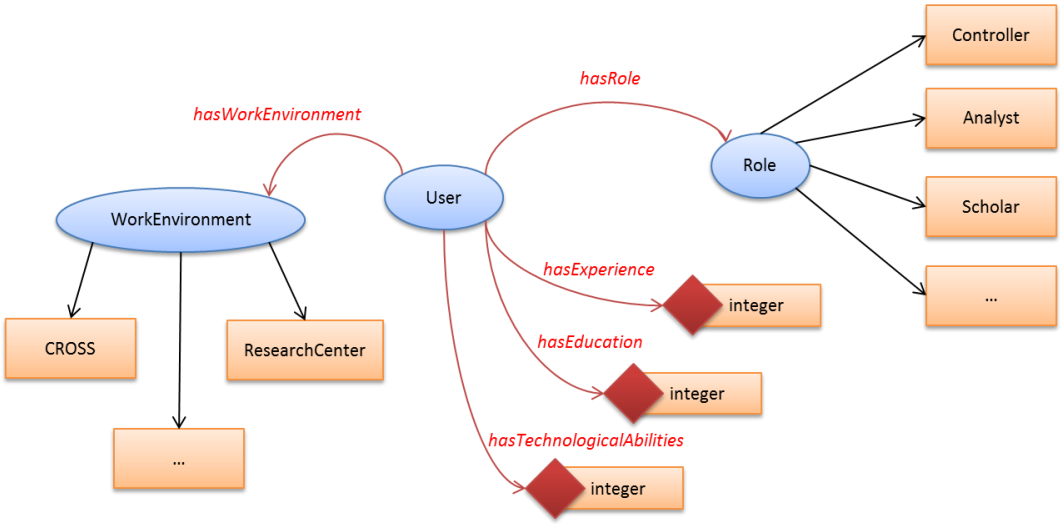


Figure 3-17. Relations sémantiques entre le concept *User* et les caractéristiques de l'utilisateur

Table 3-4. Exemples d'instances d'utilisateurs dans l'ontologie du domaine maritime

Nom instance	hasProfession	hasExperience	hasEducation	hasTechnoAbilities	hasWorkEnvironment
User_Controller	Controller	2	1	1	WE_MRCC
User_GiResearcher	Researcher	1	3	3	WE_ResearchCenter
User_Analyst	Analyst	3	3	2	WE_Office
User_Captain	Captain	2	1	1	WE_Ship

Maintenant que nous avons présenté les grands concepts qui définissent les possibles cas d'utilisation des méthodes d'analyse géovisuelle, dans le contexte de la surveillance maritime, nous présentons dans la partie suivante notre travail sur la formalisation de ces méthodes. Elles pourront ainsi être lues et comparées aux cas d'utilisation par l'environnement d'aide à l'analyse géovisuelle.

3.2.3. Modélisation des environnements d'analyse géovisuelle

Afin de caractériser une méthode de visualisation, plusieurs questions se posent : que peut-on **faire** avec cette visualisation (manipulation d'information), que peut-on **représenter** et dans quelles situations peut-on l'appliquer ? Les deux premières questions soulèvent donc des problèmes de modélisation et méta-informations : dans cette partie, nous traitons donc ces premiers points. La troisième question est, quant à elle, relative à des règles d'utilisation, que nous pourrons voir dans la partie 3.3 de ce chapitre.

3.2.3.1. Environnements et espaces de visualisation

Rappelons la définition que nous avons choisie pour caractériser un environnement d'analyse géovisuelle, selon les travaux de MacEachren (MacEachren 1995 ; MacEachren et Kraak 1997 ; MacEachren et al. 2004), Peuquet (Peuquet 1994) et Kraak (Kraak et Ormeling 1996 ; Kraak 2003a ; Kraak 2006 ; Kraak 2013). Un environnement d'analyse géovisuelle est un ensemble de méthodes de visualisation, qui permet alors l'exploration et l'analyse de grandes quantités de données dont certaines sont à composantes géographiques. Ces méthodes de visualisations peuvent être utilisées ensemble dans une même interface, afin d'avoir plusieurs vues sur les mêmes données (voir les travaux de Bédard sur les interfaces SOLAP). Les travaux de Peuquet, puis de Kraak, montrent l'importance de prendre en compte les trois aspects géographique, temporel et sémantique, dans

l'étude des données, comme nous l'avons vu dans la partie 1.3.1. Afin de modéliser un environnement d'analyse géovisuelle, la Figure 3-18 présente par un formalise UML le concept relatif à ces environnement, que l'on note ***GeoVAE*** (pour *geovisual analytics environment*), composé de différentes vues ou espaces de visualisation (Peuquet 1994 ; Heer et Agrawala 2006 ; Kraak 2006).

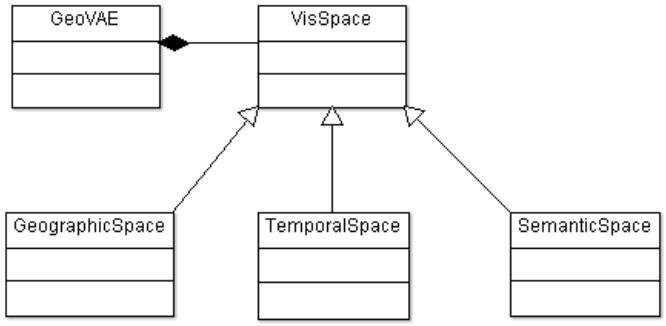


Figure 3-18. Modélisation UML pour la définition simple d'un environnement d'analyse géovisuelle

Ainsi, pour un premier niveau de classification des méthodes de visualisation d'information, nous pouvons distinguer les concepts correspondant à trois espaces, qui peuvent toutefois être complémentaires (loi de l'espace ouvert, en ontologie informatique) : ***GeographicSpace***, ***TimeSpace*** et ***SemanticSpace***. La distinction entre ces concepts se fait sur les données brutes qui peuvent être visualisées par ces méthodes : données spatiales, séries temporelles, etc. Si l'on ne précise pas que ces concepts sont distincts, les recouvrements sont possibles selon les lois des ontologies.

Nous pouvons définir ces concepts comme ayant des recouvrements, car certaines méthodes ne sont pas seulement spatiales, ou seulement temporelles par exemple : certaines, telles que le cube spatio-temporel, peuvent prendre en compte ces deux types d'informations et les mettre à un même niveau de lecture dans la présentation visuelle (axes X, Y et Z, soit les variables visuelles de position).

La Figure 3-19 présente donc ces premiers niveaux dans la caractérisation des méthodes d'analyse géovisuelle, en partant de la définition qui permet d'identifier les éléments qui nous intéressent : les espaces de visualisation (concept ***VisSpace***, dans ce cas). En suivant la définition des environnements d'analyse géovisuelle, la conceptualisation de ces environnement est donc effectuée par une relation sémantique dans l'ontologie, grâce à l'axiome ***hasVisSpace*** pour décrire un environnement (concept ***GeoVAE***). Si les différents espaces de visualisation sont dynamiquement liés au sein de l'environnement, par des techniques de sélection par exemple, l'utilisation de l'axiome ***linksVisSpace*** peut se faire, vers au moins deux méthodes. Cette même figure présente ainsi le concept général *GeoVAE* ainsi que les relations qui le relient aux espaces de visualisation.

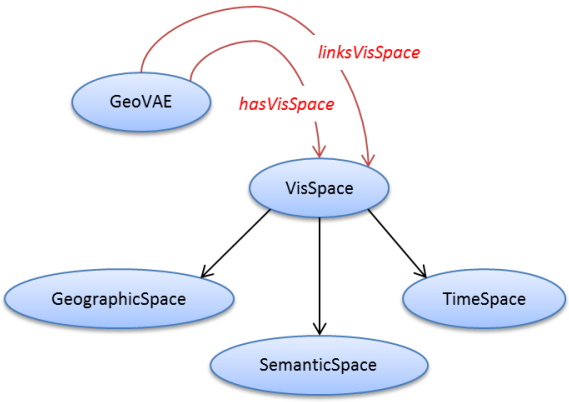


Figure 3-19. Classification générale des méthodes de visualisation d'information selon trois concepts généraux d'espaces géographiques, sémantiques et temporels

Chacune des visualisations qui composent cette taxonomie est alors une méthode qui permet de *projeter des données dites « brutes » dans un espace abstrait et visuel*, en permettant d'interagir avec l'information représentée. La sous-partie suivante décrit donc les concepts mis en place, qui permettent de décrire les différents types d'espace de visualisation des trois classes présentées ici. Une fois ces concepts de données présentés, nous décrirons la modélisation choisie pour formaliser les espaces de visualisation.

3.2.3.2. Transformation de l'information vers l'espace visuel

Dans la partie 2.2.2, nous avons présenté les travaux les plus importants menés sur modélisation du processus de visualisation, au sein de la communauté InfoVis. Nous avons vu que le *Data State Reference Model* de Chi (2000) est le plus adapté pour une approche ontologique, formelle, qui puisse être partagée et utilisée par diverses applications. Dans cette optique, nous proposons dans cette partie une modélisation du processus de visualisation de l'information (à composante spatiale, géographique) en quatre concepts liés :

- ***Data*** : la donnée brute telle qu'elle a pu être présentée dans ce chapitre (partie 3.2.2.1)
- ***DataType :*** l'abstraction analytique de la donnée, qui pourra être concrétisée par la suite dans un espace de visualisation. Ce concept s'appuie notamment sur les taxonomies de Bertin et Shneiderman (partie 2.2.1).
- ***VisualData :*** la représentation visuelle de cette information. Une « donnée visuelle » est notamment définie comme la projection d'une donnée brute, par une forme choisie. Il s'agit donc ici de la représentation du niveau *View* du DSRM. Afin de décrire ce concept, un concept supplémentaire doit intervenir.
- ***VisualShape :*** la forme visuelle de la projection de la donnée, soit le niveau d'abstraction visuelle dans le DSRM. Cela correspond notamment aux variables visuelles de Bertin, qui ont pu être présentées dans la partie 2.1.1.3.

La Figure 3-20 reprend ces différents concepts et les relations sémantiques qui les lient, formant ainsi le processus de projection de l'information vers un espace visuel, tel que décrit par Chi (2000). Sur cette figure, des cardinalités sont ajoutées aux relations qui partent de VisualData : en effet, cela nous permet de définir un concept qui « équivaut » à sa définition exacte, et ainsi limiter les erreurs possibles. En effet, un objet que nous considérons comme « données visuelles » est la combinaison d'une ou plusieurs données, représentée par une forme

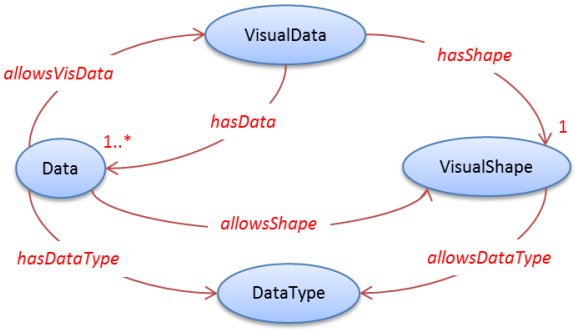


Figure 3-20. Relations sémantiques pour la description de la projection des données vers un espace visuel : concepts de données, types de données, formes visuelles et données visuelles

Afin d'illustrer les concepts de VisualData et VisualShape, nous pouvons prendre l'exemple du cube spatio-temporel, tel qu'illustré sur la Figure 3-21 : cette méthode de visualisation du mouvement représente des trajectoires dans un espace 3 dimensions, à savoir le spatial en 2D (plan vertical) et le temps en troisième dimension (axe vertical). En utilisant un terme générique, la forme visuelle de cette représentation peut être caractérisée comme une ligne en 3 dimensions (*3DLine* ou *3DGeoLine*). Les données que cette forme représente sont les données de déplacement, soit la trajectoire d'un objet et les dates correspondantes : la donnée visuelle est donc une instance du concept *VisualData*, qui relie les données *ShipTrajectory* et *ShipUTC* par une forme visuelle géoréférencée *3DGeoLine*. Cette représentation un chemin spatio-temporel, tel que l'on peut le retrouver dans la littérature (Kraak 2003b ; Kapler et Wright 2005) : dans l'ontologie, nous utilisons donc l'instance appelée *SpaceTimePath*. La Figure 3-22 représente cette instance dans l'éditeur d'ontologie Protégé.

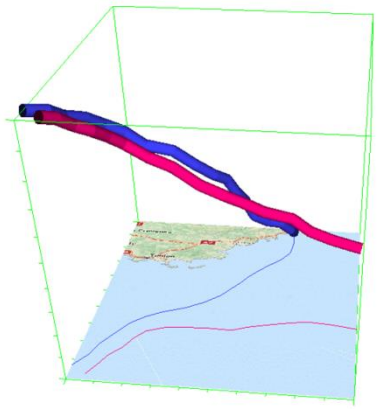


Figure 3-21. Cube spatio-temporel pour la visualisation d'une trajectoire dans l'espace et dans le temps : deux trajectoires de navires représentées dans le plan 2D et 3D

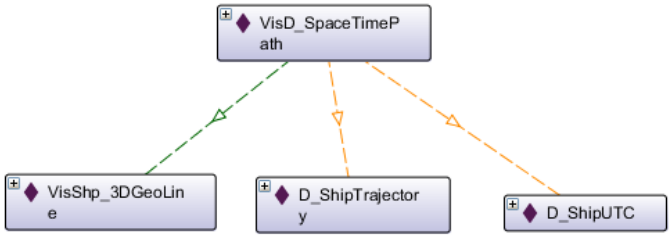


Figure 3-22. Instance du chemin spatio-temporel dans l'ontologie et liens sémantiques pour sa description

Dans un éditeur d'ontologies, ces liens et équivalence sont tels que représentés sur la Figure 3-23. L'utilisation d'axiomes pour décrire des concepts, tels que la propriété *hasDataType*, permet de définir le concept de départ comme une sous-classe d'un concept abstrait : l'ensemble de ceux qui sont reliés à DataType par cette propriété. La règle du monde ouvert stipule que ce concept n'est pas équivalent à cette définition, mais simplement inclus. Pour cela, on peut repérer sur la Figure 3-23 (a) un icône différent pour le concept de VisualData : la description de ce concept est une équivalence, telle qu'expliquée précédemment. Ainsi, toute instance de ce concept doit nécessairement respecter ces propriétés ; et toute instance qui les respecte est considérée comme appartenant à ce concept.

Afin d'avoir un catalogue de « données visuelles » qui puisse être appliqué à la caractérisation de nombreuse méthodes de visualisation, nous sommes donc partis de l'existant en visualisation d'IG pour la gestion des risques maritimes (outils opérationnels), agrémentés des visualisations proposées par la communauté GeoVis, mais moins présente dans le monde opérationnel. Nous avons ainsi pu identifier des représentations planes ou 3D, d'information statique, animée / dynamique, ou en temps réel. Les éléments les plus courants ont été introduits dans l'ontologie, sous forme d'instance, et décrit par les relations sémantiques identifiées dans cette partie.

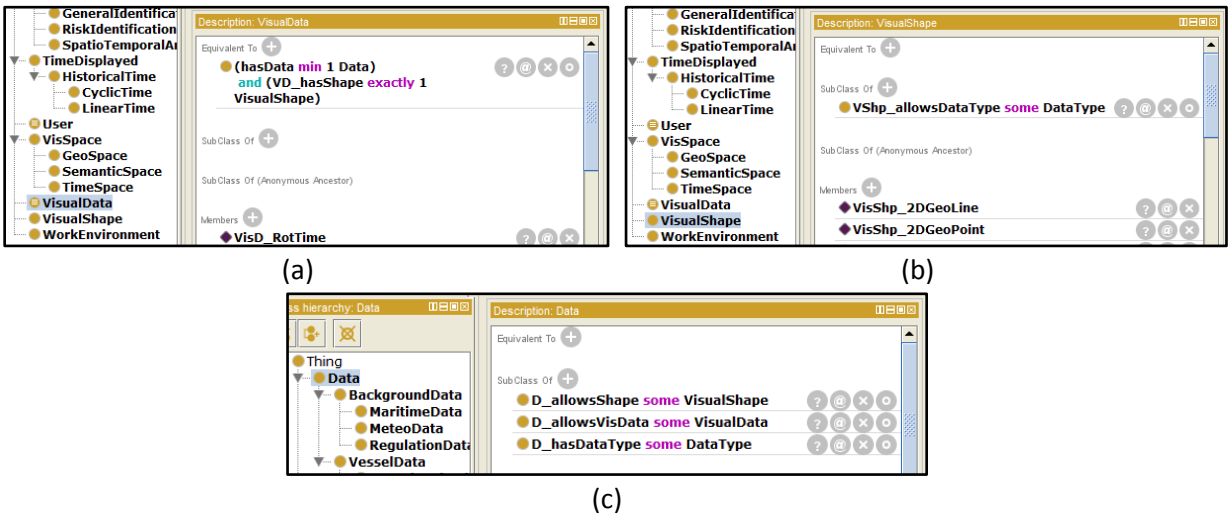


Figure 3-23. Concepts relatifs aux données et à leur projection visuelle, dans l'éditeur Protégé 4.3 : (a) les données visuelles, (b) les formes visuelles et (c) les données brutes

Une fois le processus de visualisation divisé et formalisé selon ces concepts successifs, il est alors possible de regrouper ces informations au sein de l'objet final d'espace de visualisation. La prochaine

sous-partie présente la modélisation choisie pour ce concept global, qui sert d'objet « pivot » dans ce travail de recherche.

3.2.3.3. Formalisation de l'espace de visualisation

Comme nous l'avons vu précédemment, et selon le modèle proposé par Chi (2000), nous choisissons de formaliser un espace de visualisation (*VisSpace*) comme l'utilisation d'un ensemble de données visuelles et de manipulations possibles (interactions), de façon à permettre certaines tâches d'analyse de l'information spatiale, temporelle ou sémantique. En termes d'ontologie, le concept d'espace de visualisation est alors défini par les propriétés le liant aux différentes instances *VisData* et *Task*. De plus, il est nécessaire de pouvoir renseigner les capacités et les contraintes de ces différentes méthodes en termes d'affichage de données.

De manière à caractériser les limites des méthodes de visualisations, qui peuvent être identifiées dans la littérature propre à chaque méthode de visualisation, nous utilisons les quatre propriétés suivantes :

- *displaysTime(VisSpace, TimeDisplayed)*, afin de décrire la temporalité de données que peut utiliser cette méthode. Cette méthode prend donc des valeurs de la taxonomie de la temporalité, décrite précédemment.
- *hasMaxObjects(VisSpace, positiveInteger)*, qui donne un ordre de grandeur du nombre d'objets qui peut être utilisé pour cette visualisation. De façon à simplifier la modélisation, et les règles qui en suivront, nous avons utilisé un nombre fixé de valeurs pour cette propriété : 1, 2, 10 pour des petits groupes de données, ou finalement 999 pour un très grand nombre ou un nombre « illimité » de données (par exemple, l'utilisation de cartes de densité ne limite pas le nombre d'objets).
- *isMobilityRelative(VisSpace, boolean)*, afin de décrire si la visualisation est adaptée pour un environnement mobile, qui demande donc une visualisation relative du point de vue du mobile.
- *hasAbstractionLevel(VisSpace, positiveInteger)*, de manière à indiquer le niveau d'abstraction analytique induit par cette méthode. Si les données du terrain sont directement représentées, telles que la position des navires, alors cette abstraction est nulle. Si les données proviennent d'un calcul (données « dérivées », dans la taxonomie de données) ou subissent une abstraction analytique / visuelle importante (telle que sur des cartes de Kohonen), alors cette abstraction est forte. Dans ce cas, nous donnons à ce niveau abstraction la valeur maximum de 3. Cette mesure empirique permet par la suite de mieux développer les règles d'utilisation.

La Figure 3-24 reprend donc les propriétés objets et les propriétés données qui permettent de caractériser un espace de visualisations, comme nous le proposons dans cette ontologie. Lors de la partie présentant les règles de raisonnement, nous reviendrons sur l'utilité de tels attributs.

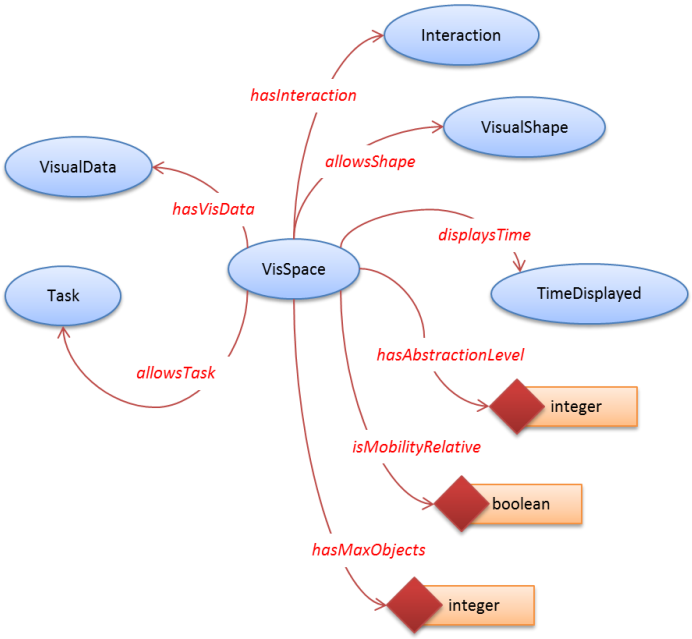


Figure 3-24. Formalisation de l'espace de visualisation par les relations entre concepts

Les variables de la visualisation et méthodes ayant été ainsi identifiées et caractérisées par ces concepts, il est nécessaire de savoir de quelle manière l'utilisateur de la visualisation peut interagir avec celles-ci. Dans la prochaine sous-partie, nous présentons donc des taxonomies de l'interaction en visualisation et leur application dans l'ontologie que nous avons développée.

3.2.3.4. Interactions proposées par la visualisation

Au cours des définitions de l'analyse géovisuelle que nous avons présentées, ou de la visualisation en général, la notion d'interaction et de manipulation de l'information est au cœur du procédé. La caractérisation des environnements d'analyse géovisuelle doit donc nécessairement prendre en compte les interactions possibles entre l'utilisateur et les données brutes, abstraites et visuelles. Nous devons donc créer un concept supplémentaire qui liste les interactions possibles au sein des méthodes d'analyse géovisuelle.

Pour la caractérisation des interactions en visualisation d'information, Roth (Roth 2012) a effectué une étude approfondie des termes les plus utilisés pour décrire l'interaction avec des interfaces de visualisation. On peut en retrouver la base dans le mantra de Shneiderman (Shneiderman 1996) pour la recherche visuelle d'information : *Overview, Zoom, Filter, Details-on-demand, Relate, History and Extract*.

Roth propose trois groupes de contrôles, qu'il divise ainsi : ***objective-based*** (orientés objectif) pour les contrôles reliés aux tâches et buts d'utilisation de la visualisation, que nous avons présentés précédemment ; **operator-based** (orientés opérateur) pour les contrôle relatifs à l'interface homme-machine que l'opérateur doit manipuler ; et finalement **operand-based** (orientés opérande) pour les transformations des propriétés visuelles des objets. Ces trois regroupements de contrôles sont effectués à partir d'un corpus d'articles scientifique spécialisés en interaction visuelle. Les figures suivantes présentent le large corps qui a été étudié, afin de proposer les familles de contrôles, ainsi que les plus importants (voir Figure 3-10, Figure 3-25 et Figure 3-26).

L'étude du groupe de contrôles suivant concerne la manipulation directe de l'information visuelle, au sein de l'environnement de visualisation. Roth reprend et compare les termes les plus employés dans la recherche en InfoVis, afin de reconnaitre les termes les plus employés et créer ce graphe de primitives. La Figure 3-25 présente le résultat de son étude des contrôles orientés opérateur. On y retrouve notamment, en premier lieu, la méthode du *brushing* ou élagage, que nous avons présenté dans la partie 2.1.3.4 du chapitre 2. La liaison entre les différentes visualisations, afin de reconnaitre l'information sous ses différents aspects, est en effet une manipulation fondamentale afin d'identifier les nombreux liens et motifs possibles dans les informations.

La Figure 3-26 représente la classification de Roth selon son terme *operand-based*. Elle est ainsi très orientée « abstraction analytique » telle que définie par Chi (Chi 2000). Cette taxonomie est donc très utile pour valider notre modèle conceptuel de données, décrit dans la partie 3.2.2 de ce chapitre. On y retrouve notamment les termes de la classification de Shneiderman, de Bertin, ainsi que la notion de temporalité.

Suite à la distinction de ces trois groupes de primitives relatives au contrôle de la visualisation, nous pouvons identifier la seconde taxonomie de Roth (primitives orientées opérateurs) comme étant la plus adaptée pour décrire les possibilités d'interaction avec l'information par la visualisation. Afin de proposer un ensemble d'interactions par les méthodes de visualisation, à prendre en compte dans l'ontologie développée, nous avons donc pris comme base cette taxonomie proposée par Roth. Nous avons ensuite pu enrichir celle-ci de termes supplémentaires à partir de cas concrets de risques en mer, qui se montrent nécessaires pour l'étude des données maritimes. Ces instances sont regroupées sous le concept général ***Interaction***, qui ne nécessite pas de subdivision pour notre étude.

Afin de décrire l'utilisation des méthodes de visualisation modélisées, nous avons de plus utilisé une relation vers l'ensemble des tâches définies (concept *Task*), nommée *allowsTask(VisSpace, Task)*. Grâce à cette propriété objet, les visualisations peuvent être caractérisées selon les tâches qu'elles permettent, telles qu'elles sont identifiées dans la littérature correspondant à chacune de ces méthodes. Cette description reprend donc, selon les termes utilisés pour décrire les buts d'utilisation et le processus de visualisation, afin de décrire les possibilités de chaque méthode de visualisation d'information.

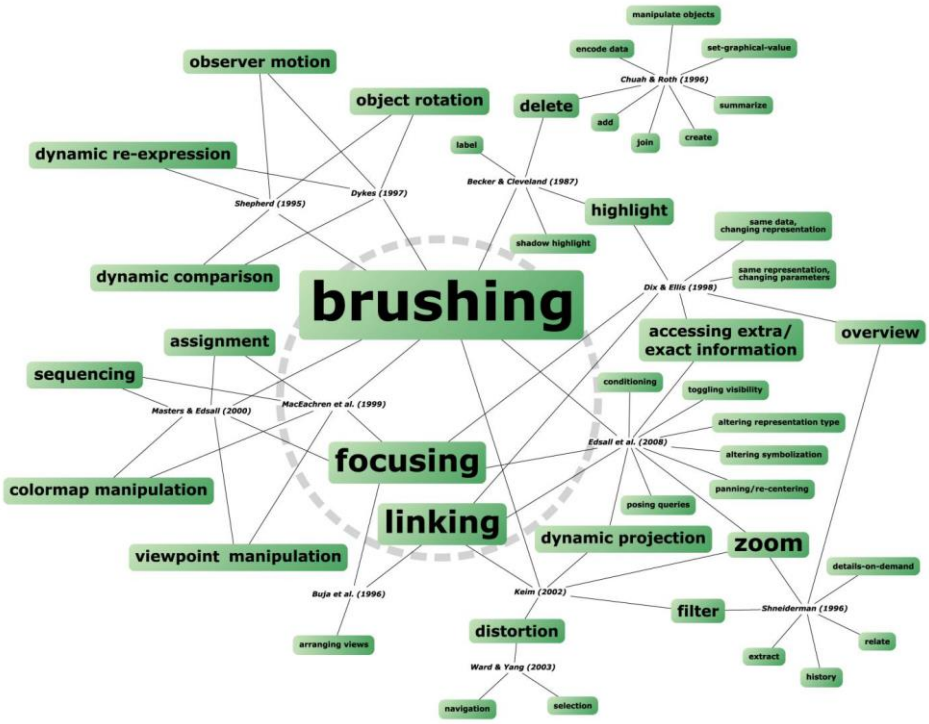


Figure 3-25. Cartographie des primitives de contrôle *orientées opérateur* (*source :* Roth 2012)

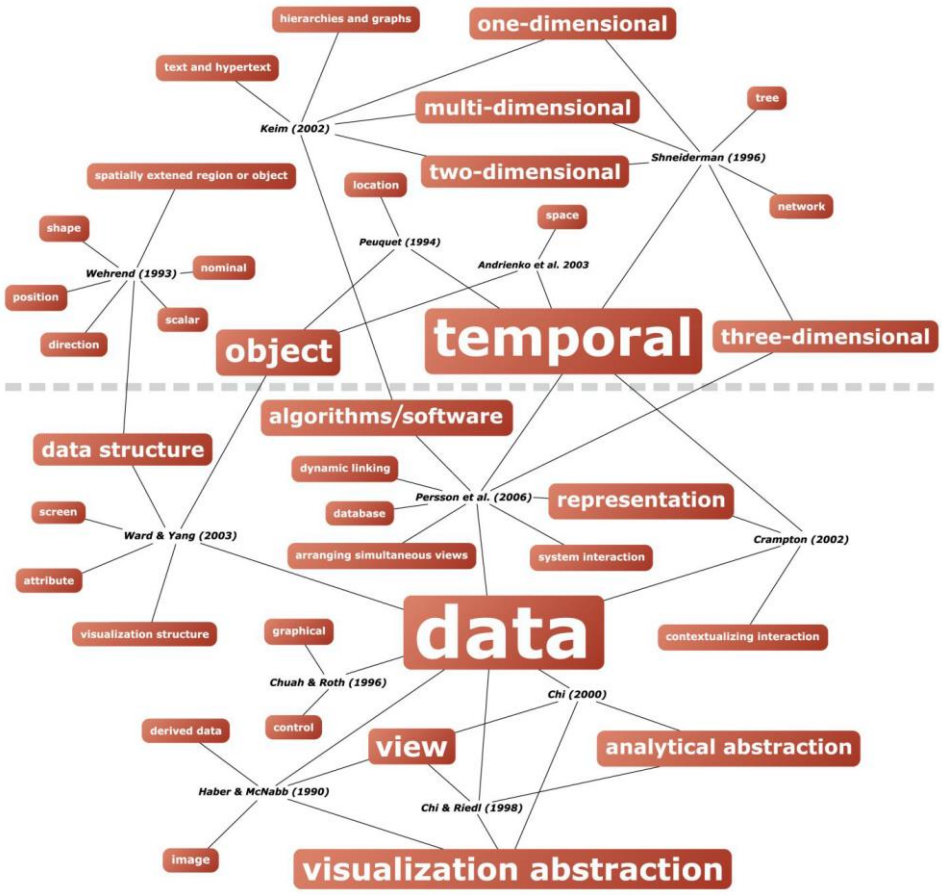


Figure 3-26. Cartographie des primitives de contrôle *orientées opérande* (*source :* Roth 2012)

3.2.4. Lien entre l'utilisateur et la visualisation

Nous avons décrit la modélisation de l'utilisateur de la visualisation, les différentes tâches qu'il peut mener par l'analyse visuelle, ainsi que la formalisation par des concepts de la visualisation d'information. Toutefois, le lien entre ces différentes ontologies doit encore être fait, de manière à comparer le cas d'utilisation et les méthodes qui peuvent être utilisées.

Pour cela, nous introduisons un nouveau concept au sein de cette ontologie, qui puisse décrire ce cas d'utilisation : le concept de *Situation*. La situation, ou bien cas d'utilisation, est utilisée comme point d'entrée pour les requêtes dans l'ontologie : ce concept abstrait décrit, pour un cas donné, l'utilisateur concerné (relation vers un profil utilisateur), le but d'utilisation ou les/la tâche(s) qu'il veut mener par la visualisation, ainsi que les données qu'il a à sa disponibilité et la temporalité de l'étude (temps réel, historique, etc.). A cela, nous ajoutons une possible information sur le risque que l'utilisateur veut analyser ou bien reconnaitre dans les données. La Figure 3-27 représente le concept Situation tel que nous l'avons défini : dans la partie gauche, sous forme de graphe, et dans la partie droite dans l'éditeur Protégé. Nous pouvons y voir les cardinalités imposées pour les relations, qui permettent de définir une équivalence pour le concept ainsi développé.

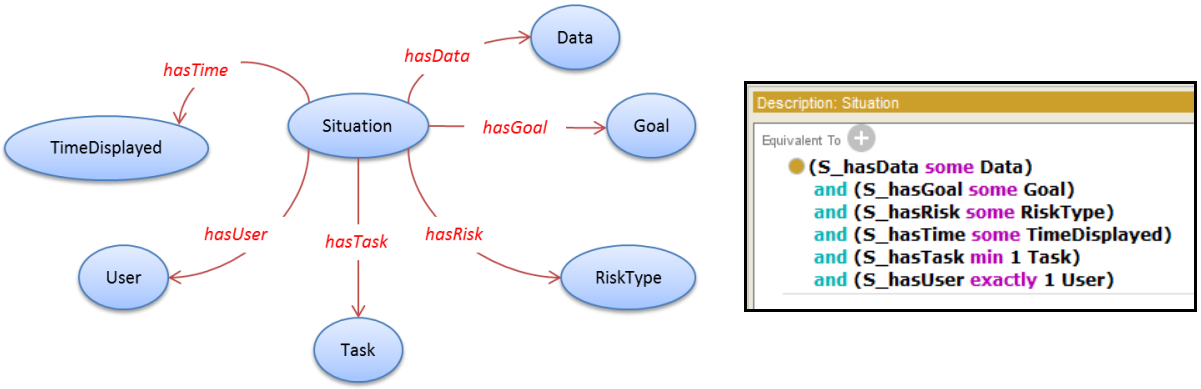


Figure 3-27. Concept Situation pour la description du cas d'utilisation et les propriétés correspondantes dans l'éditeur Protégé 4.3

L'ajout de cet objet *Situation* nous permettra ainsi de mettre en route le raisonneur, afin d'étudier le cas d'utilisation et proposer des visualisations qui soit adaptées à ce cas spécifique. Les règles qui seront alors appliquées prendront les caractéristiques de ces objets en point d'entrée de leur raisonnement. Nous reviendrons en détail sur les règles dans la partie 3.3 de ce chapitre.

3.2.5. Ontologie globale et instances développées

La Figure 3-28 reprend tous les concepts présentés dans les parties précédentes, pour présenter leurs liens par les axiomes développés. Une fois cette structure générale mise au point, les différentes instances correspondant à des objets d'études d'intérêt peuvent alors être développées. Nous présentons par la suite quelques exemples de modélisation d'utilisateurs, de scénario d'utilisation et d'environnements et méthodes d'analyse géovisuelles.

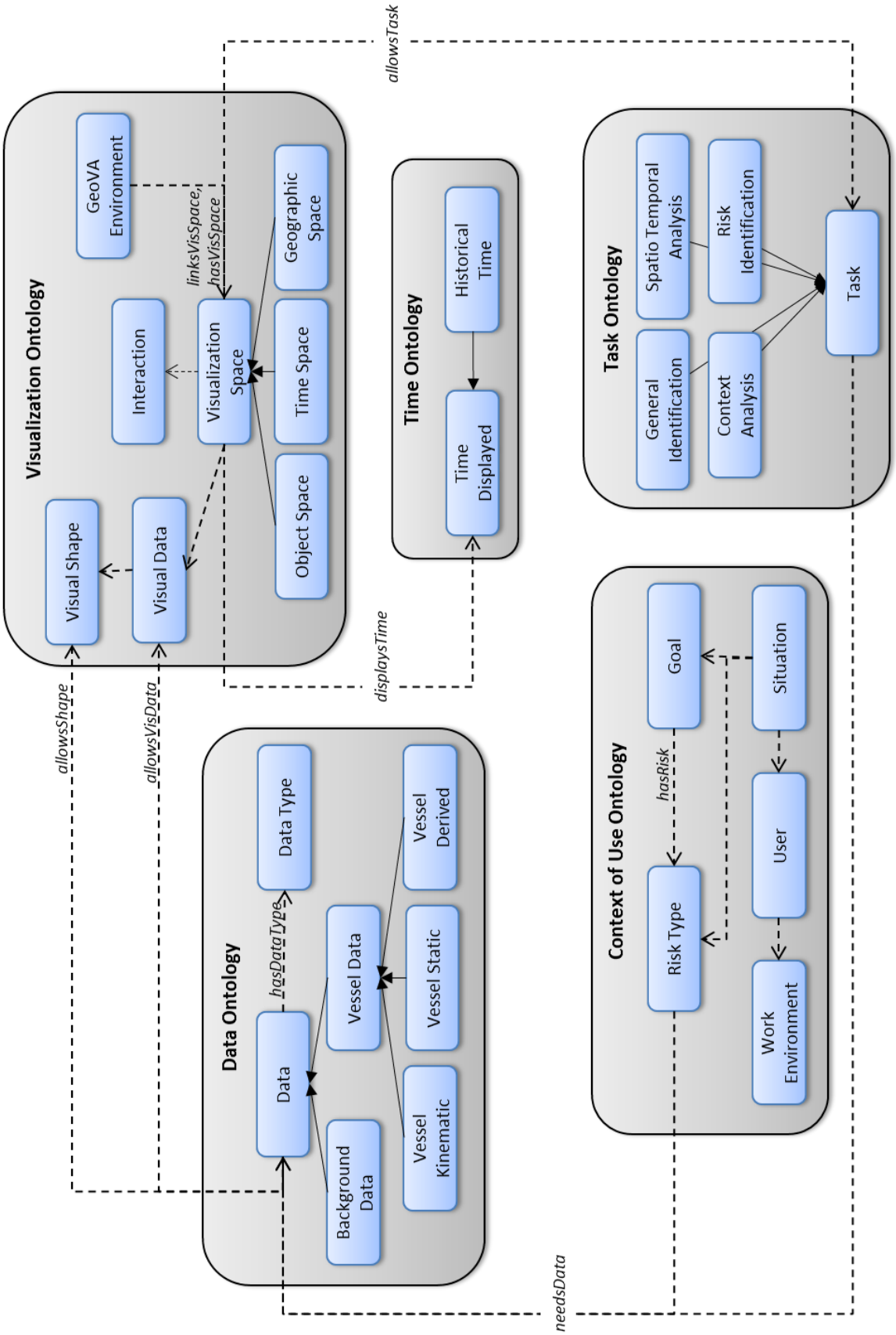
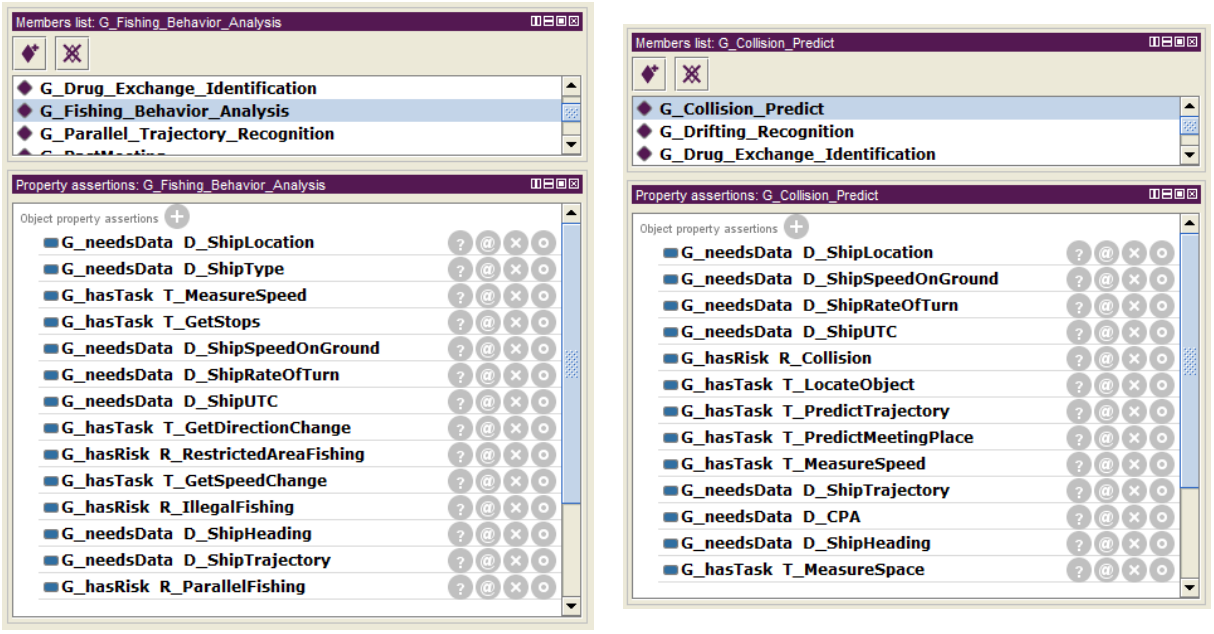


Figure 3-28. Schéma global de l'ontologie développée, selon les concepts et leurs relations hiérarchiques et sémantiques

Sur la Figure 3-29, nous présentons deux captures d'écran de l'éditeur Protégé 4.3 pour la description des étapes d'analyse pour deux scénarios test : l'analyse d'un comportement de pêche et la prédiction d'une collision, telle que nous l'avons présentée plus haut. Sur cette capture, nous pouvons voir les différentes instances du concept *Goal* (losanges violets), ainsi que les propriétés dites « objets » de l'instance sélectionnée, en bleu (relations sémantiques vers d'autres instances). Tel qu'est défini le concept *Goal*, ces instances comportent des liens vers les risques qu'elles permettent d'analyser, les données nécessaires pour mener cette étude et les tâches à mener pour cette analyse. Les tâches ne sont pas ordonnées et représentent donc le même « poids » dans la recherche des visualisations qui seront adaptées à ces études.

L'EAAG permet donc, à terme, de rechercher les méthodes et environnements d'analyse géovisuelle qui permettent de représenter ces données et effectuer ces opérations d'analyse de l'information, sans donner d'importance à l'ordre des tâches d'analyse.



(a) (b)

Figure 3-29. Formalisation des stratégies de visualisation pour deux scénarios : (a) analyse d'un comportement de pêche, (b) prédiction d'une collision

La Figure 3-30 représente la formalisation d'une méthode d'analyse géovisuelle, présentée dans une partie précédente : le cube spatio-temporel. Dans cette caractérisation de la méthode, nous utilisons une relation sémantique vers la forme visuelle de chemin spatio-temporel, tel qu'il a été décrit dans la partie 3.2.3.2 de ce chapitre. Cette caractérisation de la méthode est faite « manuellement », c'est-à-dire qu'elle est stockée dans la base de connaissances, avant l'application de règles. Afin de faire évoluer les tâches possibles via cette méthode, il est donc nécessaire de modifier le fichier de l'ontologie, par un éditeur par exemple. On retrouve sur cette figure les relations sémantiques suivantes :

- La donnée visuelle affichée
- Les types de temporalité que cette méthode permet ou non d'afficher
- Les interactions possibles
- Les tâches possibles

Sur cette figure, nous pouvons remarquer que le cube spatio-temporel appartient à la fois aux concepts d'espace de visualisation géographique et temporel : ces deux concepts n'étant pas définis comme distinct, ce cas n'est pas considéré comme une erreur dans l'ontologie. En effet, cette méthode permet à la fois de visualiser des informations spatiales, par la trajectoire des objets en mouvement, et des informations temporelles, grâce à l'utilisation de l'axe Z du cube. L'espace géographique et le temps sont donc mis sur des niveaux de lecture équivalents.

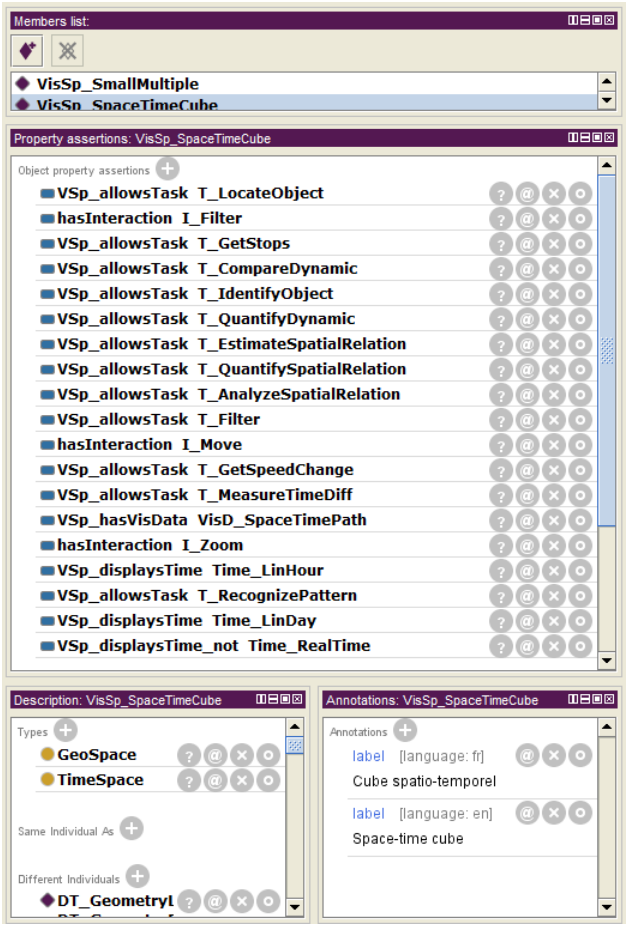


Figure 3-30. Formalisation de la méthode du cube spatio-temporel pour l'analyse d'une trajectoire dans l'espace et le temps

Afin de relier ces différentes instances, nous avons présenté le concept de *Situation*. L'utilisation des concepts liés de *Situation*, *User* et *Goal*, est alors illustrée sur la Figure 3-31, grâce aux propriétés objets. Ici, nous donnons un exemple de situation qui fait office d'informations d'entrée pour le système à base de connaissances. Les instances suivantes sont utilisées :

- Un utilisateur dont le rôle est *Controller*, qui présente une grande expérience dans le domaine maritime, mais une éducation scientifique (statistiques, mathématiques, etc.) faible et des compétences technologiques moyennes. Cela représente l'habitude de tels utilisateurs à utiliser une cartographie telle que les ECDIS, ainsi que des visualisations d'information simple telles que les diagrammes.
- Des informations traitées en temps réel (les instances de données disponibles ne sont pas indiquées et sont par défaut toutes considérées comme vraies, disponibles).

- Un but d'utilisation correspondant à l'analyse de trajectoires en vue d'une possible détection de collision. Cette instance présente les différentes tâches (non ordonnées) ainsi que les données nécessaires pour cette étude.

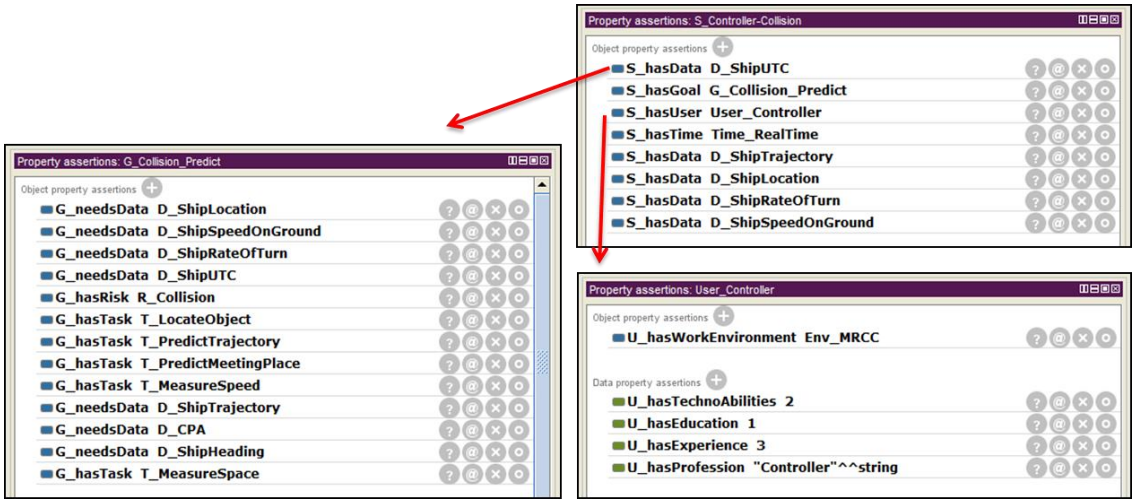


Figure 3-31. Liens sémantiques entre des instances de Situation, User et Goal

Cette situation est l'information qui permet alors de parcourir le système à base de connaissances et rechercher des méthodes de visualisation qui permettent de mener à bien ces tâches d'analyse. Pour ceci, il est encore nécessaire d'expliquer le cœur du raisonnement, c'est-à-dire l'utilisation de règles, basées sur les concepts présentés dans cette partie.

Dans la prochaine partie, nous décrivons donc les ensembles de règles que nous avons développées afin de compléter les propriétés des instances de l'ontologie et rechercher des méthodes adéquates. Afin de lire l'ensemble des méthodes conseillées, une nouvelle relation sémantique doit donc être ajoutée à l'objet central de ce raisonnement : le lien qui permet de relier la Situation aux méthodes de visualisation (*VisSpace*) et aux environnements (*GeoVAE*) qui peuvent être utilisés. Ces propriétés sont alors notées *hasVisSpace* et *hasGeoVA*. La finalité des règles développées est donc de compléter ces propriétés pour l'instance de Situation introduite dans l'ontologie.

3.3. Modélisation des règles d'inférence : recherche de solutions pour l'analyse géovisuelle

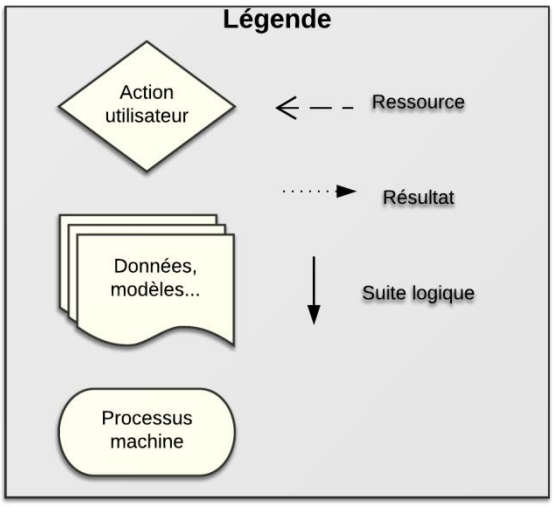
Une fois les domaines du maritime et de l'analyse géovisuelle modélisés, il est nécessaire d'ajouter des règles au sein de l'ontologie afin de réaliser une **base de connaissances** complète. Ces règles permettent d'analyser le cas d'utilisation (Situation), étudier les méthodes et les environnements modélisés, et ainsi proposer des méthodes adéquates. Ces règles viennent modifier les instances dans l'ontologie, en fonction de la validation ou non des règles. Le langage SWRL, présenté dans la partie 2.3.3 du chapitre 2, est utilisé. Nous présentons donc ici les différentes catégories de règles qui ont été intégrées dans l'ontologie, afin d'aider dans l'analyse géovisuelle de données de mouvement.

Les règles que nous présentons ici sont organisées selon plusieurs groupes logiques. Toutefois, l'ordre de leur exécution va dépendre du raisonneur utilisé. Un **raisonneur incrémental** permet en effet d'appliquer toutes les règles une à une jusqu'à ce qu'un état d'équilibre soit trouvé (toutes les première parties des règles ont été vérifiées, sans contradiction) (Grau et al. 2007). Pour cela, les

résultats des processus de raisonnements précédents sont repris pour effectuer un nouveau raisonnement. Lors de la définition des règles, nous ne pouvons donc pas imposer d'ordre logique, si ce n'est en utilisant le résultat d'une règle comme point d'entrée pour une autre.

Sur la Figure 3-32, nous présentons la place de l'intervention de l'ontologie et des règles, lors de la sélection de méthodes de visualisation, pendant l'utilisation de notre environnement d'aide à l'analyse géovisuelle. A partir des informations concernant le concept Situation, rentrées par l'utilisateur (étape 1), le choix d'une tâche d'exploration, ou de plusieurs tâches composant un but d'utilisation, l'ontologie va tout d'abord rechercher les données nécessaires. Pour cela, les descriptions ontologiques des différents concepts de buts, tâches, données, etc., sont utilisées (étape 2). Suite à cela, les règles SWRL doivent étudier la situation d'entrée et rechercher des visualisations qui correspondent à la demande (étapes 3 et 4). A la fin du processus, une liste de méthodes de visualisation doit être proposée à l'utilisateur, afin de lui laisser libre choix des méthodes qu'il souhaite utiliser pour son étude de données.

Pour cela, nous pouvons distinguer deux étapes successives dans la recherche de visualisations, qui font appel à des règles et caractéristiques différentes : une première sélection de visualisations qui permettent de représenter les données et mener les tâches demandées (sélection 1). Suite à cela, une seconde étape de sous-sélection correspond à une restriction dans les méthodes proposées : cette restriction correspond au profil de l'utilisateur et à ses contraintes d'utilisation. L'ensemble des visualisations qui en résulte (sélection 2) forme alors un groupe de méthodes qui peuvent représenter les données voulues, mener les tâches nécessaires, tout en respectant les contraintes de l'utilisateur.



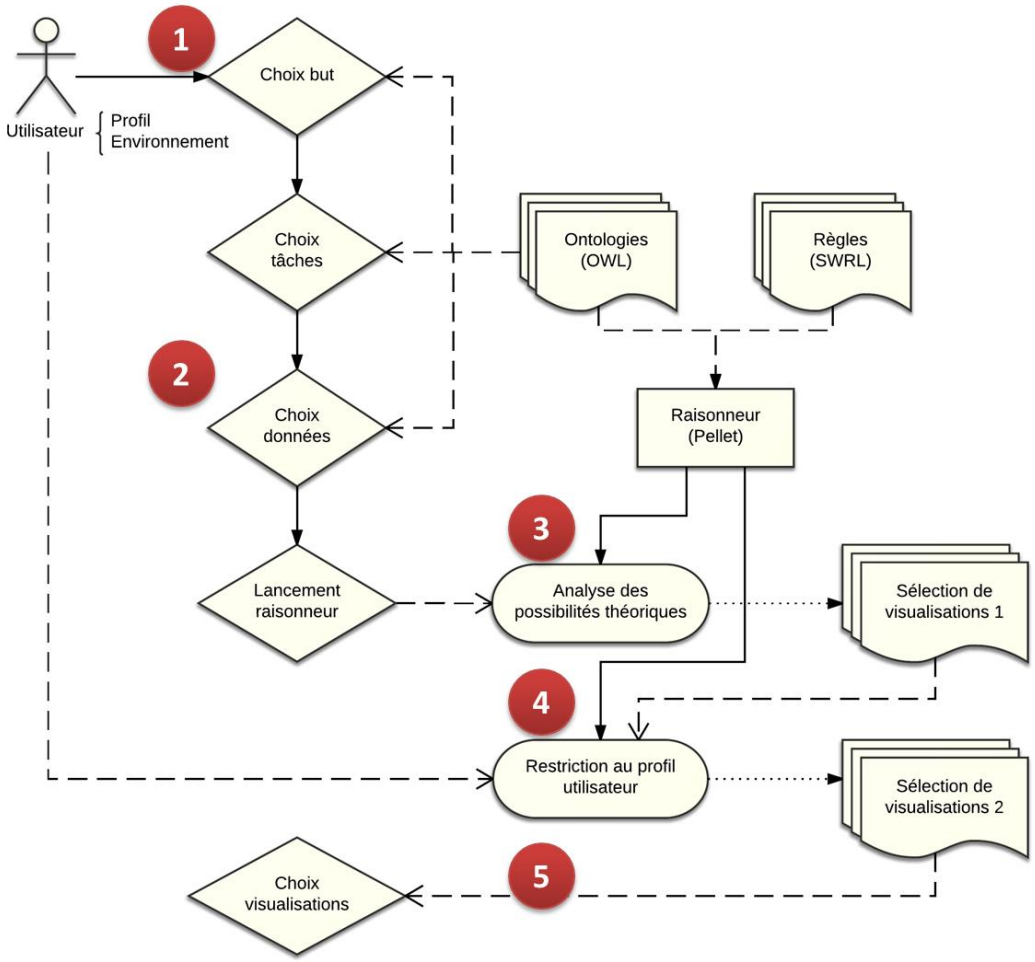


Figure 3-32. Intervention de l'ontologie et des règles pour la sélection de visualisations adaptées

Au sein de ce processus de raisonnement, nous pouvons identifier de nombreuses règles, qui pourront être appliquées de manière incrémentale, afin de sélectionner des méthodes adaptées à la situation courante. Nous pouvons identifier des groupes de règles, non ordonnés, qui correspondent aux rôles suivants :

- Compléter l'ontologie en remontant les informations des concepts de plus bas niveaux, vers les concepts les plus hauts, par la lecture des relations. Cette étape permet notamment de compléter les connaissances sur les possibilités d'utilisation de variables visuelles, si observées dans une méthode de visualisation spécifique (voir Figure 3-33).
- Ajouter des relations d'informations dites « implicitement » visualisées, bien qu'une méthode ne projette pas la donnée brute correspondante. Cela correspond à des informations qui peuvent ressortir de la visualisation, telles que la vitesse d'objets, selon la visualisation utilisée.
- Etude de la situation (cas d'utilisation) et recherche de méthodes selon leurs caractéristiques
- Catégoriser les espaces de visualisation selon les trois groupes identifiés plus haut, selon les données qu'ils permettent de représenter (espace, temps et sémantique)
- Par la redondance d'informations entre différentes visualisations, retrouver celles qui peuvent être liées dynamiquement sur plusieurs vues.

- Etudier le cas d'utilisation (Situation) en regroupant les données, tâches nécessaires, ainsi que les contraintes imposées par le profil de l'utilisateur.
- Sélectionner des visualisations et des environnements qui permettent de représenter correctement l'information nécessaire, tout en respectant les limites d'utilisation imposées par le profil de l'utilisateur.

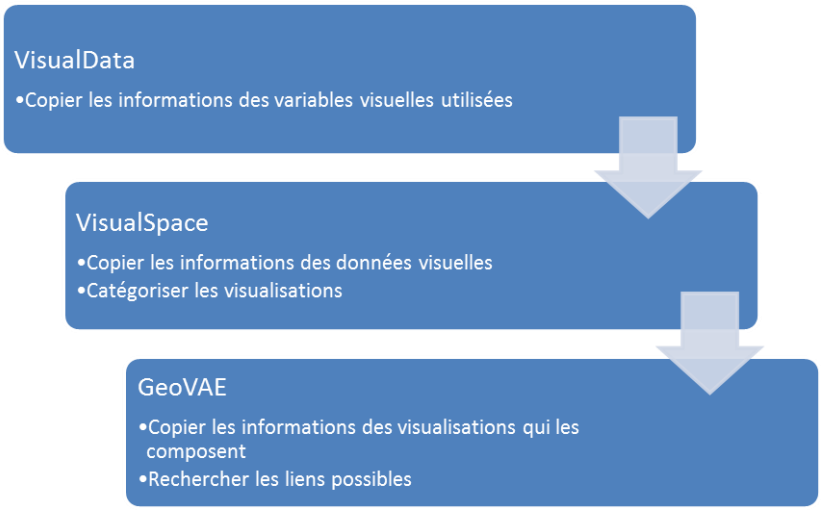


Figure 3-33. Copie des informations entre concepts, du plus bas niveau vers le plus haut

3.3.1. Plan d'exécution des règles

Les différentes règles qui peuvent être exécutées afin de sélectionner des visualisations, tel que présenté sur la Figure 3-32, viennent ajouter de nouveaux axiomes sur les concepts de Situation, VisualSpace et GeoVAE : à chaque nouvelle situation entrée dans l'ontologie, toutes les méthodes de visualisation sont comparées par rapport aux besoins identifiés, et étudiés afin de savoir si elles sont potentiellement adaptées au cas. Pour cela, de nouveaux axiomes doivent être introduits sur les concepts de VisualSpace et de GeoVAE, afin de noter leur possible adéquation au cas d'utilisation. Pour cela, nous utilisons les propriétés ***VisSp_isAdapted*** et ***GVAE_isAdapted*** afin de décrire, de manière générale, si un espace (méthode) de visualisation ou un environnement d'analyse géovisuelle sont adaptés à la situation (valeur booléenne). Ces attributs dérivent de plusieurs attributs, à valeurs booléennes, qui permettent de décrire de façon plus précise l'adéquation de ces méthodes aux différents critères de la situation : le profil de l'utilisateur, les données et leur temporalité, les tâches à effectuer, etc.

Afin de contextualiser l'ensemble des règles qui sont développées, la Figure 3-34 présente un plan des nombreuses étapes à effectuer lors de l'analyse de la situation, des données, des tâches, etc. Contrairement à la figure, constituée de blocs successifs, les règles doivent être pensées sans ordre dans leur exécution.

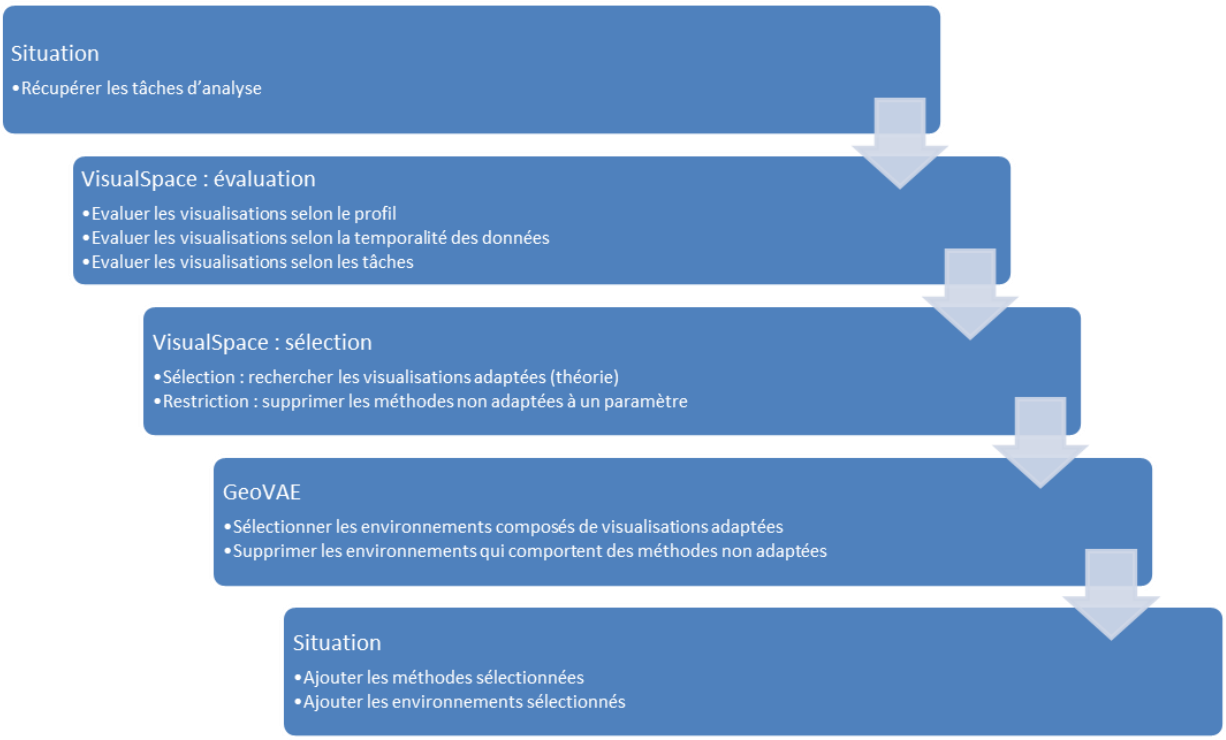


Figure 3-34. Suite logique pour la recherche de visualisations adaptées : parcours des différents concepts

La sous-partie suivante présente donc les différents groupes de règles qui ont été mis au point afin d'opérer ces différentes étapes proposées ici. Bien que le schéma indique des flèches, qui suivent un raisonnement tel que nous pouvons le faire, les règles sont exécutées de manière simultanée, jusqu'à un état d'équilibre de ces observations, c'est-à-dire une fois que toutes les règles sont validées, et qu'il n'existe pas de contradiction au sein de l'ontologie.

3.3.2. Ensemble de règles développées : processus de raisonnement

Les premières règles développées permettent de remonter les informations des instances dans la hiérarchie des concepts, tel que nous l'avons expliqué précédemment. Cela permet, entre autre, de simplifier la lecture de l'ontologie lors des règles suivantes, en simplifiant les règles. Ces propriétés sont donc « copiées » selon les liens d'héritage et les relations entre concepts créés.

Pour plus de lisibilité dans ces règles, nous posons le résultat de la règle sur une ligne séparée du reste des prédicats. Sur chaque propriété, les premières lettres correspondent au concept auquel celles-ci s'appliquent, pour plus de clarté dans l'ontologie et les règles. Sur la Figure 3-35, nous rappelons le formalisme du langage SWRL, tel que nous l'avons utilisé pour représenter les règles suivantes.

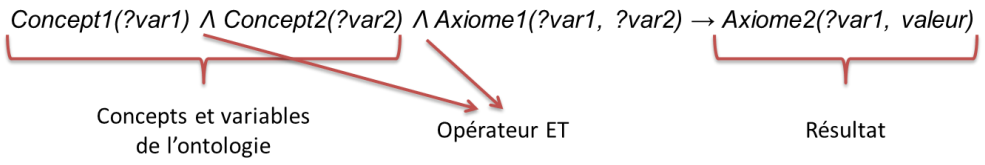


Figure 3-35. Formalisme SWRL utilisé

3.3.2.1. Remontée des caractéristiques des données vers les espaces de visualisations et les environnements d'analyse géovisuelle

Cette première règle permet de « remonter » ou copier les propriétés des concepts de formes et de données visuelles, vers le concept d'espace de visualisation. Selon les attributs de la forme (Shape) qu'utilise une donnée visuelle (VisualData), cette dernière va donc adopter les dimensions que la forme : par exemple, plane ou 3D. La règle correspondante, en SWRL, est alors donnée dans le cadre suivant.

VisualData(?vd)
Λ VD_hasShape(?vd, ?sh)
Λ VShp_hasVisDimension(?sh, ?dim)
→ VD_hasVisDimension(?vd, ?dim)

Cette propriété remonte ensuite vers les instances des espaces de visualisation (VisSpace), ainsi que les instances des environnements d'analyse géovisuelle, de la même façon. La règle suivante donne l'exemple pour le concept d'espace de visualisation.

VisSpace(?vs)
Λ VSp_hasVisData(?vs, ?vd)
Λ VD_hasVisDimension(?vd, ?dim)
→ VSp_hasVisDimension(?vs, ?dim)

De manière similaire, les caractéristiques des espaces de visualisation, telles que les interactions proposées, les données visuelles utilisées, les tâches possibles, etc., sont ensuite remontées vers la description des environnements d'analyse géovisuelles, selon le lien de composition. Les règles utilisées pour cela sont proches de la règle précédente.

3.3.2.2. Données autorisées selon les espaces de visualisation

Il est important de pouvoir savoir, sans faire de requête complexe et en lisant l'ontologie, quelles sont les données qu'une méthode de visualisation peut représenter. Cela se retrouve en redescendant le processus de visualisation, en partant d'une visualisation vers les données de base. De la même façon que précédemment, les informations relatives aux données brutes auxquelles correspondent les données visuelles peuvent donc être copiées au niveau des instances des espaces de visualisation.

VisSpace(?vs)
Λ VSp_hasVisData(?vs, ?vd)
Λ hasData(?vd, ?data)
→ hasData(?vs, ?data)

Afin de compléter les connaissances sur les différentes méthodes de visualisation, la règle suivante permet de renseigner les informations relatives aux possibilités de représentations, qui n'auraient pas été renseignées dans l'ontologie. Cette règle permet de savoir quels sont les types de données que permettent de représenter les différentes méthodes de visualisation, selon les données visuelles qu'elles comportent. Ainsi, cette information peut être utilisée par la suite afin de renseigner sur l'utilisation possible d'une visualisation, sur d'autres données du même type.

VisSpace(?vs)
Λ VSp_hasVisData(?vs, ?vd)
Λ VD_hasShape(?vd, ?sh)
Λ VShp_allowsDataType(?sh, ?dt)
→ VSp_allowsDataType(?vs, ?dt)

3.3.2.3. Informations dites implicites

Certaines données ne sont pas utilisées de façon « directe » par certaines méthodes de visualisation : la valeur brute n'est pas projetée et représentée par une variable visuelle, mais la valeur peut être déduite d'autres informations représentées, et être représentée de manière implicite par la visualisation (Dörk et al. 2012). Dans leurs travaux, Dörk et al. identifient plusieurs types d'informations implicites qui peuvent être retrouvées par la visualisation : des relations multiples, des motifs invisibles, ou encore des informations additionnelles.

En termes de données spatio-temporelles, l'exemple du cube spatio-temporel peut être repris pour illustrer ce principe : lors de la représentation d'un chemin spatio-temporel, la donnée de vitesse n'est pas utilisée. Toutefois, le fait de représenter à la fois le déplacement dans l'espace géographique plan et dans le temps, par l'axe Z, permet de déduire des informations de vitesse (Zhong et al. 2010) : les lieux et moments d'accélération, d'arrêt, etc. Au sein de cette méthode, la vitesse peut donc être considérée comme information implicite, mais qui doit toutefois être considérée comme visible par cette méthode.

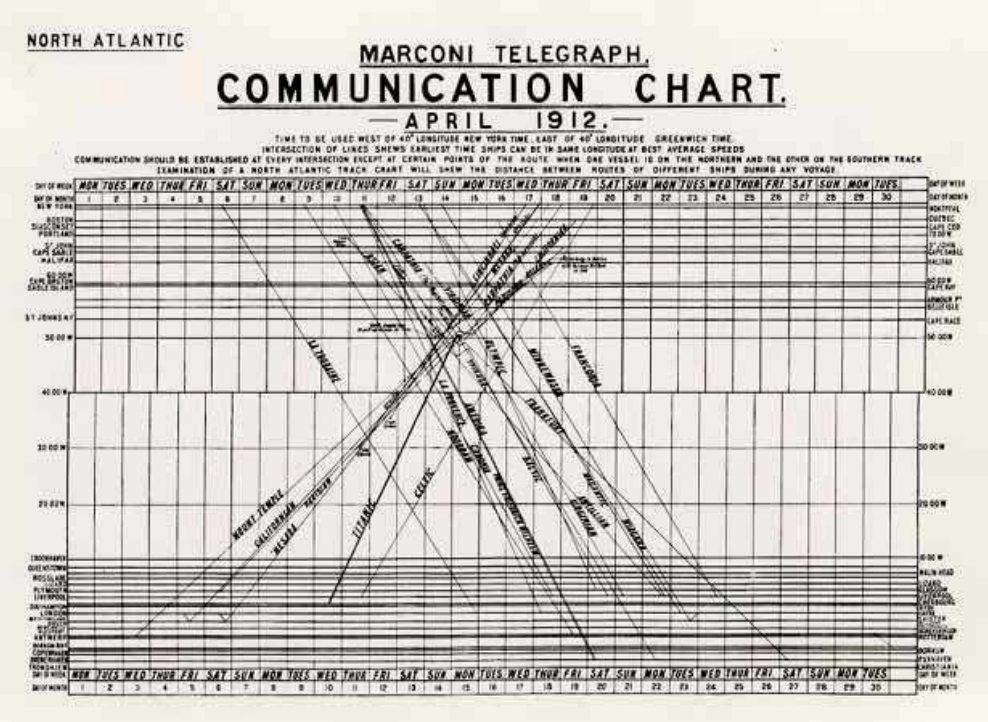


Figure 3-36. Diagramme de communication maritime du Titanic, qui permet d'identifier les postes terrestres et les navires avec lesquels communiquer pendant la traversée

La Figure 3-36 présente un autre type de représentation d'information spatiale et temporelle, beaucoup plus ancienne mais tout aussi efficace. Il s'agit du diagramme de communication élaboré pour la traversée de l'Atlantique par le Titanic (en 1912), par le physicien Marconi (1847-1937). Ce

diagramme permet de visualiser la présence des navires dans l'espace et dans le temps, le long des longitudes. Cette représentation des voyages donne les lieux de départ et d'arrivée, ainsi que le jour de départ et d'arrivée. Il permet d'estimer, dans cette projection spécifique, les vitesses de croisière ainsi que la proximité possible entre plusieurs navires. Ces informations ne sont pas représentées directement sur ce diagramme, mais sont implicitement visualisées.

Plusieurs règles peuvent alors être déduites afin d'ajouter ces descriptions supplémentaires sur les instances de visualisation (a fortiori, sur les environnements d'analyse géovisuelle aussi). Sans avoir de donnée visuelle qui représente exactement une donnée brute, il est important de pouvoir renseigner celle-ci comme information représentée. On retrouve cette notion de données implicite dans les lois de la Gestalt, grâce à la loi de la forme (un objet non entièrement représenté, mais implicitement reconnu). La règle suivante représente la donnée implicite de la vitesse d'un objet, reconnue dans le cas d'un cube spatio-temporel.

VisSpace(?vs)
Λ VSp_hasVisData(?vs, ?vd)
Λ hasData(?vd, D_ShipTrajectory)
Λ hasData(?vd, D_ShipUTC)
Λ VSp_hasVisDimension(?vs, 3)
→ hasData(?vs, D_ShipSpeedOnGround)

3.3.2.4. Liens possibles

Dans le domaine de l'analyse géovisuelle, la redondance d'informations sur plusieurs vues utilisées en parallèle est une notion capitale, comme nous l'avons vu dans les parties précédentes. En effet, représenter une même information de plusieurs manières, c'est-à-dire grâce à plusieurs espaces de visualisation, permet de distinguer des groupes de données sur certaines vues. C'est pour cela que des interactions telles que le *highlighting*, le *brushing*, etc., ont pu être proposés. Afin de prendre en compte cette notion de redondance de l'information dans des vues liées, au sein d'un environnement d'analyse géovisuelle, une règle permet de trouver les visualisations qui peuvent être liées. Pour cela, celles-ci doivent appartenir à un même environnement, et représenter des données communes. La règles suivante peut alors être formulée, en utilisant la propriété *linksVis* du concept GeoVAE (environnement d'analyse géovisuelle).

GeoVAE(?gva)
Λ GVAE_hasVisSpace(?gva, ?vs1)
Λ GVAE_hasVisSpace(?gva, ?vs2)
Λ hasData(?vs1, ?d1)
Λ hasData(?vs2, ?d2)
Λ SameAs (?d1, ?d2)
Λ DifferentFrom (?vs1, ?vs2)
→ GVAE_linksVis(?gva, ?vs1) Λ GVAE_linksVis(?gva, ?vs2)

Les premières règles présentées ici correspondent à une première phase, lors du raisonnement, pour compléter les informations des données visuelles, des espaces de visualisation, ou des environnements d'analyse géovisuelle. Ces « mises à jour » des nombreuses instances sont utilisées afin de compléter les descriptions de chaque instance, plutôt que de remonter les hiérarchies et les liens lors des autres règles à développer. Ainsi, les instances modélisées dans la base de

connaissances sont complétées par ces règles, qui renseignent des propriétés supplémentaires pour décrire le processus d'analyse géovisuelle, ainsi que les utilisations possibles des différents concepts.

Les règles suivantes portent donc sur l'étude du cas d'utilisation, et la sélection de visualisations qui soient adaptées à cette situation. Nous présentons donc les différentes étapes de sélection et restriction, telles qu'introduites sur la Figure 3-32.

3.3.2.5. Etude de la situation

Comme nous l'avons présenté au début de cette partie, le raisonnement dans l'ontologie part d'une nouvelle instance Situation, unique au sein de l'ontologie. A partir des caractéristiques de cet objet, relatives à l'utilisateur et aux tâches qu'il veut effectuer, les visualisations et les environnements doivent être étudiés à partir de leurs caractéristiques modélisées. Pour cela, plusieurs caractéristiques booléennes sont ajoutées aux visualisations, de manière à rechercher les méthodes adaptées. Ces caractéristiques booléennes sont les suivantes : *isAdapted-Data*, *isAdapted-Education*, *isAdapted-Experience*, *isAdapted-Task*, *isAdapted-Techno* et *isAdapted-Time*. Au final, une unique propriété permet de faire la décision, en combinant ces informations : *isAdapted*.

3.3.2.5.1. Règles sur la temporalité de données

Lors du développement de règles en SWRL, il n'est pas possible d'utiliser des négations. Cela vient de la règle du monde ouvert des ontologies, contrairement aux bases de données classiques. En effet, si une instance n'a pas une propriété clairement définie, cela ne signifie pas qu'elle ne la vérifie pas. Cette règle de monde ouvert, qui bloque les négations de propriétés, a été une contrainte importante lors du développement des concepts et règles, notamment pour l'étude de la temporalité des données.

Afin de faire face à cette limite, nous avons dû introduire une propriété supplémentaire, qui soit contraire à la propriété *displaysTime* des espaces de visualisation. Cette propriété (positive) décrit en effet les types de temporalité qui peuvent être affichées par une visualisation, telle que des données temps réel. A l'inverse, nous avons ajouté une propriété booléenne *displaysTime_not*, qui permet de compléter la description des visualisations, et simplifie les règles à développer. Cette propriété permet donc de renseigner quant aux types de temporalité non supportés par des méthodes de visualisation.

A partir de ces deux propriétés, positives et négatives, nous avons mis au point les règles suivantes qui décrivent l'utilisation possible ou non de la visualisation, selon la temporalité des données choisies. Cette première règle décrit les cas « positifs » :

Situation(?s)
∧ VisSpace(?vp)
∧ S_hasTime(?s, ?t1)
∧ VSp_displaysTime_not(?vp, ?t2)
∧ SameAs (?t1, ?t2)
→ VSp_isAdapted-Time(?vp, false)

L'inverse de cette règle est alors obtenu en utilisant les propriétés contraires :

Situation(?s)
Λ VisSpace(?vp)
Λ S_hasTime(?s, ?t1)
Λ VSp_displaysTime(?vp, ?t2)
Λ SameAs (?t1, ?t2)
→ VSp_isAdapted-Time(?vp, true)

Cette règle aboutit donc à l'ajout d'une caractéristique supplémentaire sur chaque espace de visualisation modélisé dans l'ontologie, tel que le montre la capture d'écran sur la Figure 3-37.

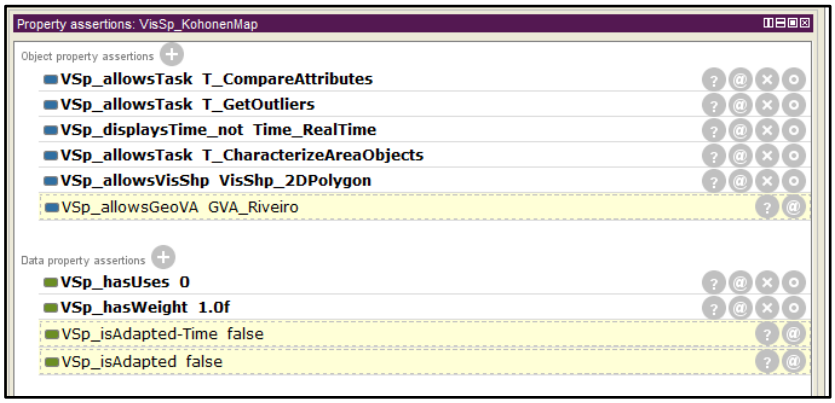


Figure 3-37. Visualisation non adaptée à la temporalité des données : étude de données temps réel par une visualisation type carte de Kohonen

3.3.2.5.2. Règle sur l'adaptation aux tâches

La règle centrale pour la recherche de méthodes dépend ensuite des tâches d'analyse qui doivent être menées, afin de répondre aux questions de l'utilisateur. Pour cela, les tâches nécessaires qui caractérisent l'instance de Situation sont comparées aux tâches admises par les différentes méthodes de visualisation. Le principe de cette règle est alors de trouver les méthodes qui proposent les tâches recherchées. Une méthode peut proposer bien plus de tâche, et sera tout de même prise en compte : cette première sélection de méthodes recherche toutes les possibilités (voir Figure 3-38).

Pour cela, la règle correspondante est la suivante, qui permet de valider le booléen de la caractéristique *isAdapted-Task*, pour toutes les visualisations au sein de l'ontologie. Les méthodes sélectionnées sont donc celles qui proposent au moins une des tâches nécessaires à l'exploration ou l'analyse à mener.

Situation(?s)
Λ VisSpace(?vp)
Λ S_hasTask(?s, ?t1)
Λ allowsTask(?vp, ?t2)
Λ SameAs (?t1, ?t2)
→ VSp_isAdapted-Task(?vp, true)

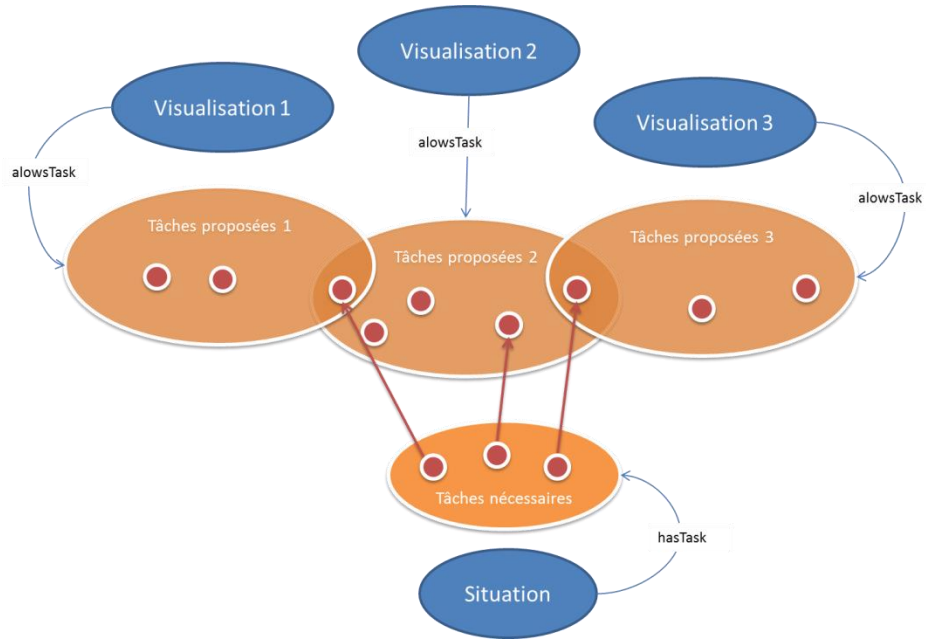


Figure 3-38. Sélection de méthodes par les tâches proposées

Au terme de ces deux premières étapes, toutes les méthodes de visualisation sont alors caractérisées comme adaptées ou non à la temporalité des données, et aux tâches à effectuer. Le monde ouvert de l'ontologie ne permet pas de caractériser les méthodes qui ne proposent pas de tâches nécessaires : ces visualisations n'ont donc pas de valeur pour la propriété *isAdapted-Task*.

Les règles suivantes concernent alors la restriction de cette première sélection de visualisations, selon les contraintes de l'utilisateur et des données qu'il a à sa disposition.

3.3.2.5.3. Restrictions selon le profil de l'utilisateur

La propriété correspondant aux compétences technologiques de l'utilisateur sont ici illustrées pour le cas d'utilisation de visualisation en 3 dimensions (cube spatio-temporel, mur de trajectoires, etc.). A titre empirique, nous proposons une règle qui permette l'utilisation de ce type de visualisations pour les utilisateurs les plus expérimentés en termes de technologie ; au contraire des visualisations utilisées dans les centres opérationnels, qui sont des représentations planes de l'espace ou de diagrammes. Un tel type d'utilisateur peut représenter une personne plus apte à utiliser des visualisations complexes et modernes, telles que dans le domaine de la recherche en visualisation / information géographique.

Nous proposons plusieurs règles pour la gestion du nombre de dimensions de l'espace visuel utilisé. La première règle donne par défaut la valeur *True* à cet attribut, si la dimension visuelle est supérieure ou égale à 1 (espace linéaire).

Situation(?s)
Λ VisSpace(?vs)
Λ S_hasUser(?s, ?u)
Λ U_hasTechnoAbilities(?u, ?tA)
Λ VSp_hasVisDimension(?vs, ?dim)
Λ greaterThanOrEqual(?tA, 1)
Λ lessThan(?dim, 3)
→ VSp_isAdapted-Techno(?vs, true)

Pour des visualisations à 3 dimensions, deux règles complémentaires doivent être utilisées pour paramétrer cet attribut comme vrai ou faux, selon les compétences technologiques. Suivant l'échelle à trois niveaux que nous utilisons, les règles suivantes peuvent donc être utilisées :

Situation(?s)
Λ VisSpace(?vs)
Λ S_hasUser(?s, ?u)
Λ U_hasTechnoAbilities(?u, ?tA)
Λ VSp_hasVisDimension(?vs, ?dim)
Λ equal(?dim, 3) Λ lessThan(?tA, 3)
→ VSp_isAdapted-Techno(?vs, false)

Situation(?s)
Λ VisSpace(?vs)
Λ S_hasUser(?s, ?u)
Λ U_hasTechnoAbilities(?u, ?tA)
Λ VSp_hasVisDimension(?vs, ?dim)
Λ equal(?dim, 3) Λ equal(?tA, 3)
→ VSp_isAdapted-Techno(?vs, true)

Pour chaque instance de visualisation, l'attribut *isAdapted* doit indiquer si cette méthode est globalement adaptée à la situation, en prenant en compte toutes les contraintes. Cet attribut se trouve donc être la combinaison de toutes les autres caractéristiques identifiées dans cette partie, selon la formule :

$$isAdapted = isAdapted_Data \times isAdapted_MaxObjects \times isAdapted_MinObjects \times isAdapted_Education \times isAdapted_Experience \times isAdapted_Task \times isAdapted_Techno \times isAdapted_Time \times isAdapted_WorkEnv$$

En termes de règle, cela revient à dire que la valeur de la propriété *isAdapted* est Faux si au moins l'un des facteurs est Faux. Autrement, on peut considérer sa valeur comme Vrai. Toutefois, la règle du monde ouvert des ontologies ne permet pas de considérer si une méthode est totalement adaptée à chacun de ces attributs ; sauf si l'attribut est Faux. Cette limite d'utilisation des règles nous a donc imposé de revoir le mode de sélection des visualisations adaptées. Parmi les facteurs de cette équation, ceux qui limitent la création d'une règle finale sont *isAdapted-Data* et *isAdapted-Task* : leur valeurs dépendent des propriétés que l'instance de visualisation comporte.

En effet, si une visualisation n'est pas renseignée sur les données qu'elle peut représenter, leur relation n'est toutefois pas considérée comme inexistante. Une règle qui lirait les données reliées à une visualisation la considérerait donc comme adaptée aux données, car il n'est pas expressivement dit qu'il n'existe pas de liens entre ces instances. Le problème existait avec la variable de temporalité, mais nous avons mis au point deux paramètres inverses, que sont *displaysTime* et *displaysTime_not*. Etant donné le grand catalogue des données et des tâches disponibles, il n'est pas possible de considérer une telle relation entre les concepts VisSpace et Data.

Afin de remédier à cette limite, la valeur *isAdapted* ne prend donc pas en compte ces paramètres, mais ne considère que les facteurs qui ont nécessairement une valeur Vrai ou Faux. La formule se réécrit donc :

$$isAdapted = isAdapted_MaxObjects \times isAdapted_MinObjects \times isAdapted_Education \times isAdapted_Experience \times isAdapted_Task \times isAdapted_Techno \times isAdapted_Time \times isAdapted_WorkEnv$$

Une fois les visualisations adaptées identifiées, les règles finales permettent de les relier à la situation, afin de pouvoir les proposer à l'utilisateur, suite à sa requête.

3.3.2.5.4. Sélection finale des visualisations et des environnements

Afin de pallier aux limites des règles que nous avons identifiées précédemment, la sélection finale de visualisation se fait via deux règles supplémentaires. Pour cela, nous recherchons les méthodes qui vérifient le paramètre *isAdapted*, et qui de plus permettent de représenter au moins une des données désirées, et de réaliser au moins une des tâches d'analyse identifiées. Si ces paramètres sont vérifiés, la visualisation peut alors être reliée à l'instance de Situation, grâce à la propriété *hasVisSpace* (voir Figure 3-39). La règle correspondant à cette étape est alors la suivante :

Situation(?S)
Λ VisSpace(?vs)
Λ S_hasData(?S, ?d)
Λ hasData(?vs, ?d)
Λ S_hasTask(?S, ?t)
Λ allowsTask(?vs, ?t)
Λ VSp_isAdapted(?vs, true)
→ S_hasVisSpace(?S, ?vs)

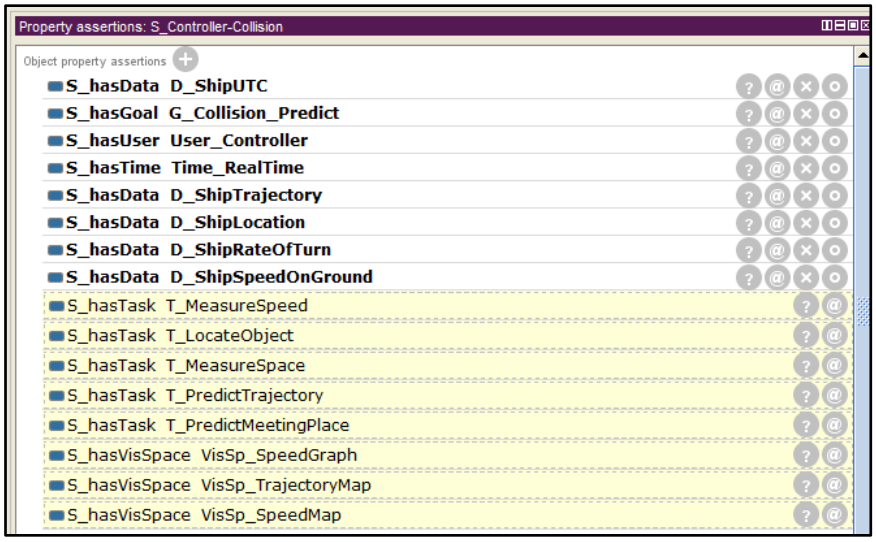


Figure 3-39. Liste des espaces de visualisation adaptés à une situation données

De manière similaire, si un environnement d'analyse géovisuelle est composé d'au moins une visualisation qui a été sélectionnée pour le cas d'utilisation, et si cet environnement ne comporte pas de visualisations qui ne soit pas adaptée, alors il peut être proposé comme environnement possible dans ce cas. La règle pour ceci correspond aux formulations SWRL suivante :

GeoVAE(?gva)
Λ VisSpace(?vs)
Λ GVAE_hasVisSpace(?gva, ?vs)
Λ VSp_isAdapted(?vs, false)
→ GVAE_isAdapted(?s, false)

Situation(?s)
Λ GeoVAE(?gva)
Λ VisSpace(?vs)
Λ S_hasVisSpace(?s, ?vs)
Λ GVAE_isAdapted(?gva, true)
→ S_hasGeoVAE(?s, ?gva)

Conclusion du Chapitre 3

Dans ce chapitre, nous avons défini le terme nouveau d'environnement d'aide à l'analyse géovisuelle, qui s'inscrit dans la continuité des systèmes d'aide à la décision et des environnements d'aide à la résolution de problèmes. Un EAAG est donc un système informatique fondé sur une formalisation de la connaissance en visualisation d'information géographique, en complément de la formalisation de connaissances pour un domaine d'application donné. Afin de présenter la réalisation d'un tel environnement, nous avons choisi le domaine de l'étude des mouvements dans le trafic maritime, pour l'étude des risques. Nous avons présenté le développement de ce système en portant notre choix sur l'utilisation d'ontologies, dans une optique de partage de la connaissance.

La deuxième partie de ce chapitre nous a donc permis de présenter l'ensemble des concepts couverts par cette ontologie, afin de décrire formellement le processus de visualisation, l'analyse de données de mouvement par la visualisation, les risques maritimes et les informations qui décrivent le cas d'utilisation. L'ensemble de ces concepts, des relations qui les lient et des instances qui les composent, permet ainsi de définir une base de connaissances pour la démonstration de cet environnement d'aide à l'analyse géovisuelle.

Dans la troisième partie du chapitre, nous avons présenté la dernière composante de cette modélisation que forme l'ensemble des règles de fonctionnement. Ces règles forment une base pour le raisonnement, à partir des concepts de l'ontologie, selon des connaissances théoriques (le fonctionnement de la visualisation) et des faits concrets pour le domaine d'application. Cet ensemble de règle permet alors d'étudier l'ensemble des concepts, l'instance d'un cas d'utilisation en entrée, les différentes visualisations modélisées, afin de sélectionner un sous-ensemble de visualisations qui puissent être adaptées, utilisées et intégrées au sein d'un environnement d'analyse géovisuelle complet.

L'ensemble de ce modèle a été présenté lors d'une conférence internationale sur le domaine de la visualisation d'information (Vatin et Napoli 2014). Dans le chapitre suivant, nous présentons notre preuve de concept pour l'utilisation d'un EAAG. Pour cela, une interface visuelle pour l'utilisation de ce système à base de connaissances est proposée. L'outil web FishEye est introduit, afin de présenter des exemples concrets de visualisation du trafic maritime pour son exploration et son analyse. Nous expliquons ainsi le lien entre le démonstrateur de l'EAAG et FishEye. Pour cela, nous allons étudier un cas concret de risque maritime, et détailler les différentes étapes d'analyse et d'interprétation de ces données.

CHAPITRE 4 - EXEMPLE D'UTILISATION DE L'ENVIRONNEMENT D'AIDE A L'ANALYSE GEOVISUELLE

Le chapitre précédent nous a permis d'introduire le concept d'environnement d'aide à l'analyse géovisuelle, par une définition générale et un exemple de développement fondé sur une base de connaissances et des règles. Nous avons présenté les nombreux concepts nécessaires pour définir un tel environnement, les relations entre ses concepts et le fonctionnement d'un raisonneur pour analyser un cas d'utilisation et proposer des visualisations qui soient adaptées au contexte d'étude.

Dans ce chapitre, nous illustrons l'utilisation d'un tel environnement d'aide à l'analyse géovisuelle, et nous présentons plusieurs cas d'analyse du trafic maritime pour la reconnaissance de risques. Pour cela, nous présentons le système développé au sein du CRC pour la visualisation et l'analyse des données maritimes, nommé FishEye. Nous décrivons le rôle et l'architecture de cet outil orienté web, ainsi que ses fonctionnalités principales. Par la suite, nous présentons le prototype d'utilisation de l'environnement d'aide à l'analyse géovisuelle, permettant de faire le lien entre la base de connaissances et le système de visualisation développé. Grâce à ces différents outils, nous introduisons alors un scénario d'intérêt pour l'étude de données de trafic maritime, et nous menons une exploration par la visualisation. Nous concluons ce chapitre par une discussion sur cette approche pour guider dans l'analyse géovisuelle, et son utilisation dans le domaine maritime.

4.1. Plateforme web de géovisualisation maritime : FishEye

4.1.1. Rôle de la plateforme

Un portail cartographique maritime (Lagesse 2011) a été développé au sein de l'équipe de recherche. Cet outil, appelé FishEye, a eu pour premiers buts d'intégrer les données utiles pour les travaux de recherche du Centre de recherche sur les Risques et les Crises, ainsi que de servir de catalogue d'informations maritimes et de visualisations de positions en temps réel des navires. Par la suite, cet outil a évolué afin de devenir une plateforme de tests pour les outils issus des différents travaux de recherche du CRC sur le thème de la sécurité et sûreté de l'espace maritime : plusieurs modules développés par les chercheurs pour la détection et l'analyse ont été intégrés, ainsi que l'utilisation de règles pour l'analyse des données géographiques.

L'utilisation de l'outil FishEye a donc évolué d'un simple outil d'affichage du trafic maritime, tel qu'il était en 2011, vers un outil complet pour l'aide à l'analyse de données de mouvement. L'expertise sur la visualisation d'information spatio-temporelles a notamment été intégrée, et les différents aspects relatifs à l'analyse de données passées ont été mis en place. Les principales fonctionnalités de FishEye sont les suivantes :

- Visualiser et analyser les trajectoires des navires, qui sont enregistrées dans des bases de données pour une analyse a posteriori
- Afficher les données règlementaires de l'espace maritime, dérivée des données ENC au format vecteur
- Comparer les comportements des navires par rapport à des modèles de risques formalisés et lever des alertes (Vandecasteele 2012 ; Vandecasteele et al. 2014)
- Etudier la dangerosité des zones de trafic maritime par l'étude d'accidents passés ou de zones de fréquentation élevée
- Etudier la fréquentation de zones d'intérêt, telles que les ports, les zones sensibles, les canaux et détroits, etc.
- Simuler des comportements tels que la dérive d'un objet, ou la collision de deux navires
- Rejouer l'état d'une situation passée dans une zone choisie
- Etudier les comportements des navires grâce à la fouille de données, afin d'extraire des comportements dangereux, anormaux, etc. (Idiri 2013)

Afin de pouvoir intégrer ces nombreux travaux de recherche, une grande partie de ceux-ci ont été développés selon une architecture web, ou grâce à des langages informatiques adaptés à une telle utilisation (par exemple, des applications Java). Pour cela, FishEye se fonde sur des librairies web en partie open-source (Ext, GeoExt, OpenLayers), afin d'utiliser les données vecteur et raster d'un serveur géographique local (GeoServer). Ce serveur géographique est aussi utilisé pour plusieurs projets de recherche sur la surveillance maritime, pour l'UE et l'ANR.

Les données en entrée de cette application sont les données AIS des navires, que nos partenaires industriels nous fournissent en temps réel dans le cadre de projets R&D. Ces données sont principalement situées sur la Méditerranée et la Manche, qui sont des enjeux majoritaires pour la sécurité et la sûreté du trafic maritime européen. En supplément de ces informations dynamiques, cet outils utilise des données descriptives du contexte maritime (météorologie, océanographie, etc.) et la réglementation dérivée des ENC. La trajectoire des navires, leur vitesse, les arrêts peuvent être ainsi visualisés, ainsi que d'autres attributs diffusés par les AIS. Des données de météorologie sont intégrées à ce système, et mises à jour toutes les 6 heures. A cela, nous pouvons de plus ajouter des informations dérivées des processus de data-mining et de reconnaissances automatique de zones dangereuses, ou de comportements anormaux.

La Figure 4-1 représente une architecture N tiers, telle que nous l'avons adoptée afin de développer FishEye : les sources d'informations proviennent de fichiers de données et de styles, de bases de données, et d'autres serveurs en cascade ; ces nombreuses sources sont alors centralisées sur un serveur géographique. Cette plateforme utilisant des données sensibles (positions des navires, informations des navires, etc.), nous avons choisi de ne la proposer qu'en interne à notre laboratoire.

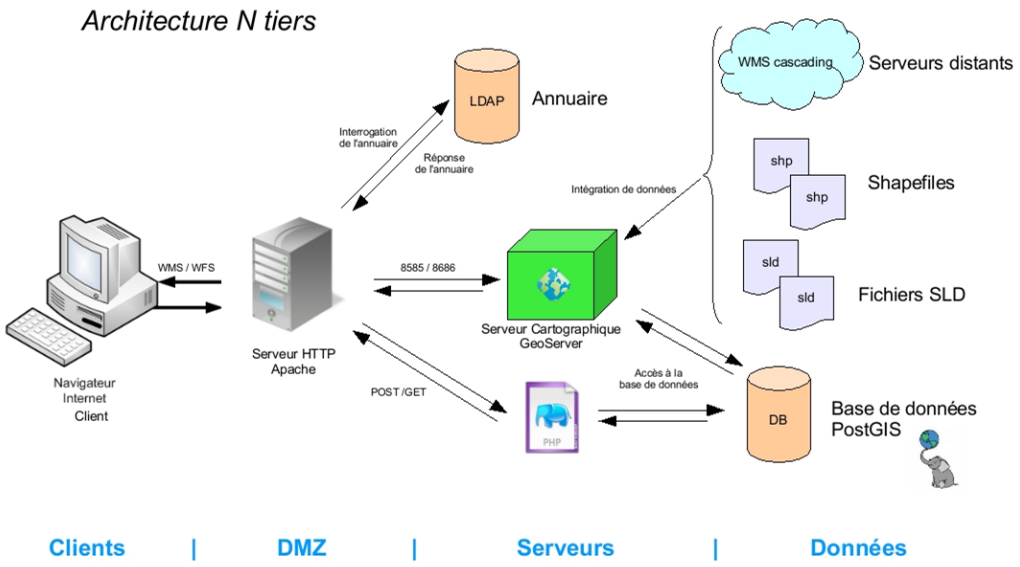


Figure 4-1. Architecture d'une application N tiers pour la gestion d'informations géographiques

Dans la sous-partie suivante, nous présentons quelques exemples de visualisations d'informations et de fonctionnalités de cet outil, pour l'étude des données de trafic maritime.

4.1.2. Exemples de visualisations

De nombreuses fonctionnalités de visualisation et d'analyse des mouvements de navires ont été intégrées dans la plateforme FishEye. Ces fonctionnalités sont accessibles depuis une barre d'outils (horizontale) ou bien par un clic droit sur un objet d'intérêt (navire, port, etc.). Une liste de couches permet de classer toutes les informations du serveur géographique selon leur thématique, et leur échelle (mondiale, ou européenne). La Figure 4-2 présente les nombreuses fonctionnalités de cette barre d'outils. Ces outils peuvent être utilisés pour ajouter / dessiner des données sur la carte, faire des mesures (espace, temps), simuler des trajectoires ou dérives, filtrer les données affichées, visualiser des informations de manière plus complexe, etc.

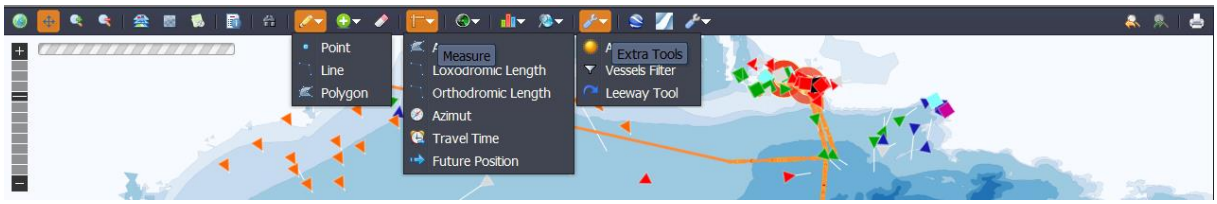


Figure 4-2. Barre d'outils de FishEye et sous-menus correspondants

La Figure 4-3 présente la visualisation d'une trajectoire de tanker (triangle rouge) sur 48h, où chaque point orange correspond à un message AIS reçu (coordonnées, vitesse, cap, etc.). La qualité du signal radio émis peut provoquer de longs espaces dans l'espace et le temps sans positions successives proches, ce qui explique des trajectoires parfois simplifiées (long segment rectiligne et absence de données). La Figure 4-4 reprend cette même trajectoire, mais avec une représentation de la dynamique de ce navire ainsi que des informations sur son identité : nous visualisons donc la vitesse le long de sa trajectoire passée, en plus du graphique de vitesse, ainsi que les arrêts détectés (vitesse

nulle ou inférieure à une certaine valeur). Sur cette figure, il s'agit d'un profil de vitesse et une forme de trajectoire régulière d'un navire tanker.

La Figure 4-5 illustre la pénétration d'un navire dans une zone sensible. Grâce à un système d'analyse automatisé, basé sur des ontologies spatio-temporelle (Vandecasteele et Napoli 2012), les comportements anormaux ou dangereux peuvent être détectés et listé dans une fenêtre d'information. Enfin, la Figure 4-6 illustre une requête sur la fréquentation d'une zone définie : le nombre de navires ayant traversé une zone tampon autour du port de Toulon sur une durée de deux semaines, qui est alors affiché sous forme de diagramme à bâtons. Un autre onglet de cette fenêtre permet de consulter ces informations sous forme de table, notamment avec le temps passé dans cette zone par chaque navire.

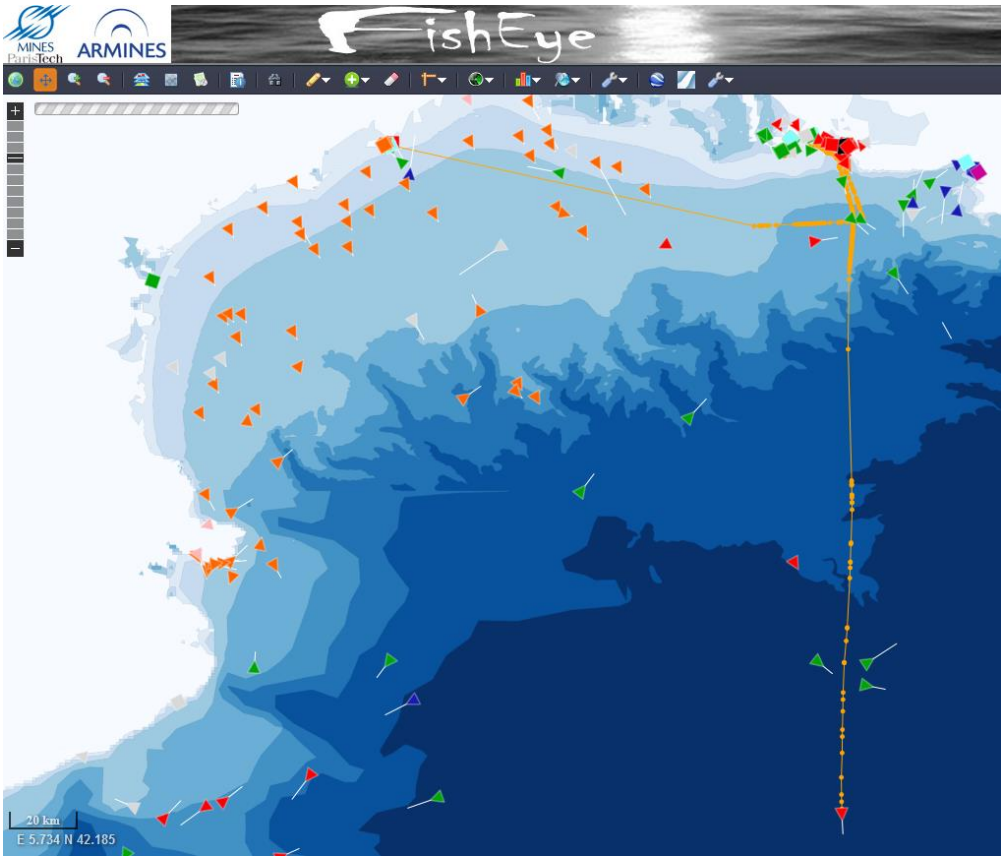


Figure 4-3. Visualisation de la trajectoire d'un tanker en Mer Méditerranée

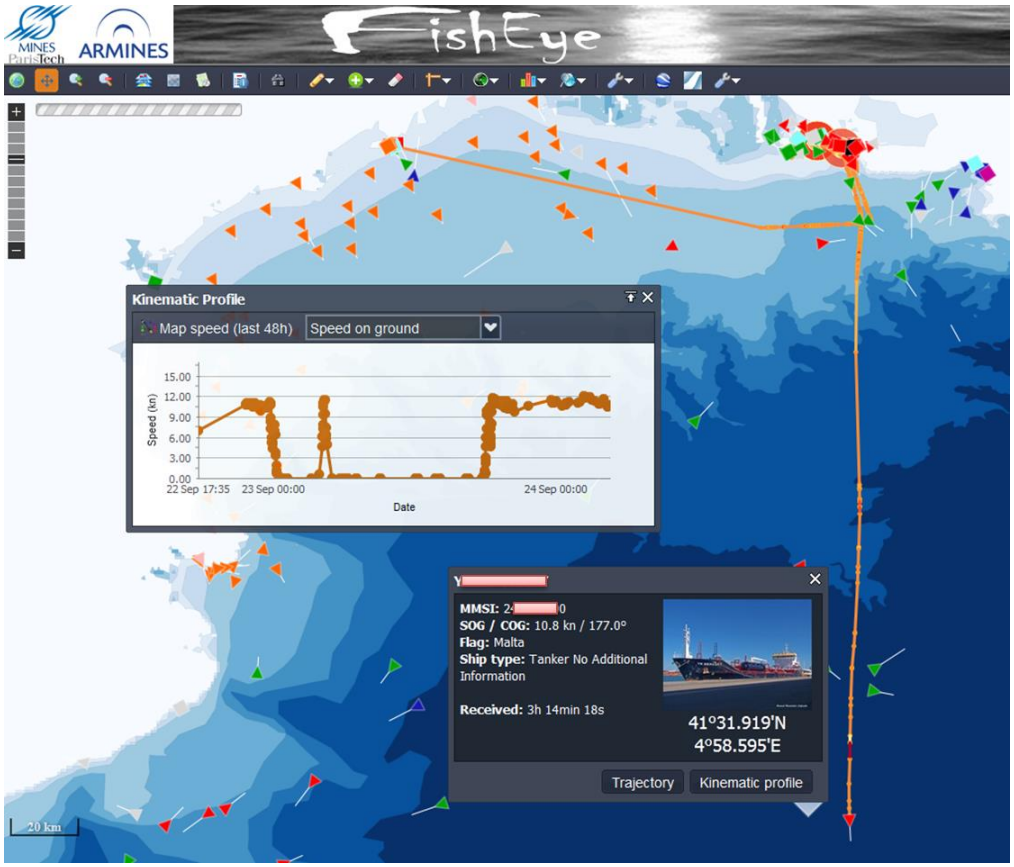


Figure 4-4. Visualisation de la trajectoire, vitesse, arrêts et informations sur un tanker

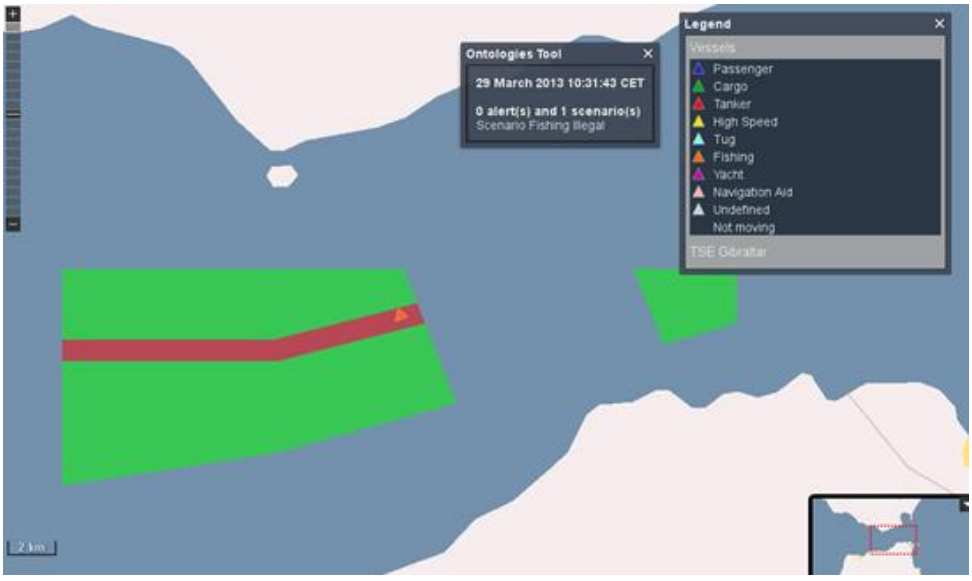


Figure 4-5. Reconnaissance d'un comportement anormal par raisonnement ontologique

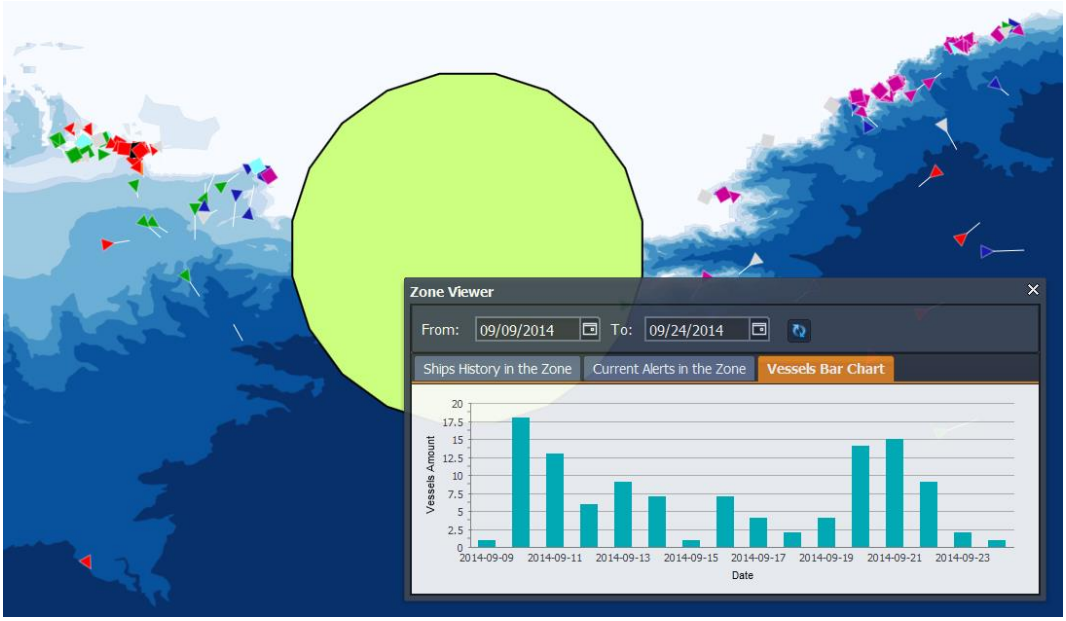


Figure 4-6. Fréquentation d'une zone tampon autour du port de Toulon, pendant deux semaines

L'amélioration de cette plateforme, d'un outil d'affichage vers un outil d'analyse, et les essais de visualisations ont été le cœur de notre travail au cours de la première année de cette thèse. Au cours de cette période, nous avons testé plusieurs méthodes de visualisations d'information géographique, appliquées au domaine maritime. Par la suite, notre travail s'est centré sur la formalisation de ces méthodes qui permettent l'analyse du trafic maritime, par des ontologies, tel que nous l'avons décrit dans le chapitre 2 de ce manuscrit. La prochaine sous-partie présente le prototype d'utilisation des connaissances ainsi formalisées, que nous avons alors intégré au sein de FishEye.

4.1.3. Interface d'utilisation de l'environnement d'aide à l'analyse géovisuelle

Comme nous l'avons présenté précédemment, l'interface entre un système à base de connaissances (système expert, PSE, SAD, etc.) est aussi importante que les connaissances elles-mêmes : c'est en effet cette IHM (interface homme-machine) qui permet la communication entre la formalisation et l'utilisateur. Elle doit donc être claire, simple d'utilisation et intuitive (Gallopoulos et al. 1992 ; Jackson 1999), afin de ne pas être un premier frein pour la future utilisation des solutions proposées. Le but de l'interface relative à l'EAAG que nous proposons doit donc servir de module de recherche de visualisations adaptées et de sélection : elle permet donc de lire différents objets de l'ontologie et d'en introduire de nouveaux, afin de lancer un raisonnement propre au cas d'utilisation. La Figure 4-7 présente les cas d'utilisation d'une telle interface par un diagramme UML : ces cas sont basés sur les concepts développés dans l'ontologie, qui définissent une situation (profil d'utilisateur, but d'utilisation et tâches, données, etc).

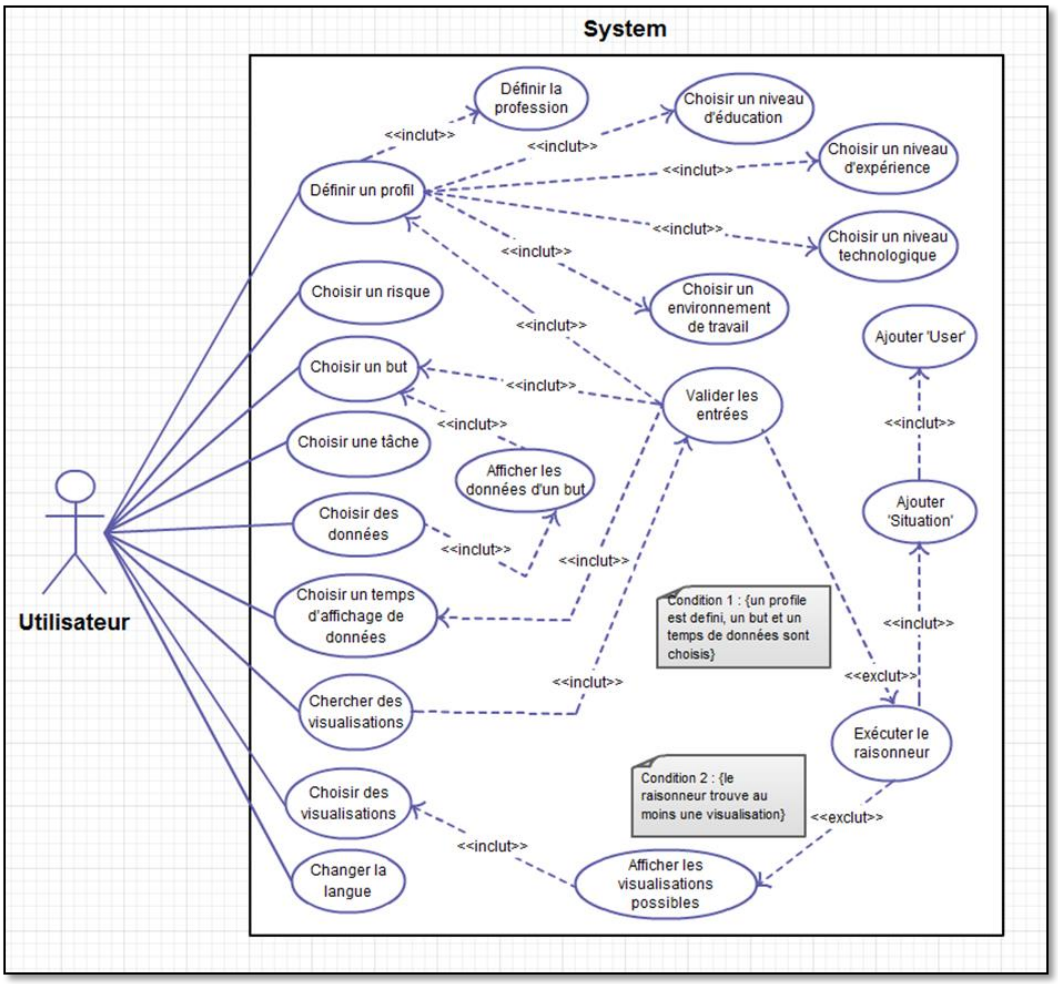


Figure 4-7. Cas d'utilisation pour l'interface entre l'utilisateur et l'ontologie au sein d'un environnement d'aide à l'analyse géovisuelle

De manière à pouvoir exploiter un tel prototype au sein de la plateforme web FishEye, nous avons choisi de développer un outil qui soit utilisable à la fois de manière indépendante (*standalone*) ou bien sur le web, par un appel de fonction. La prochaine sous-partie présente donc les choix techniques pour le développement d'une telle interface pour l'utilisation d'un système à base de connaissances, fondés sur les ontologies.

4.1.3.1. Développement du module de sélection

4.1.3.1.1. Une architecture Web

Nous avons fait le choix d'utiliser une **architecture client-serveur** pour le développement d'une telle application, afin de pouvoir facilement gérer l'ontologie côté client par des mises à jours régulières, ainsi que de permettre une utilisation locale depuis le terminal d'un client pour lire les concepts et ajouter de nouveaux objets propres à la situation. En effet, les informations relatives à la situation, ou cas d'utilisation, ne doivent pas être écrites de manière permanente : il est alors plus simple de récupérer l'ontologie « à distance » dans un fichier temporaire et local à chaque exécution, puis de modifier ce fichier selon la requête de l'utilisateur. Afin d'enregistrer les détails de chaque exécution, ces données sont ensuite envoyées sur le serveur, qui exécute alors un script pour stocker celles-ci. Pour une telle architecture, trois composantes principales ont été identifiées :

- Le client informatique, qui est amené à utiliser le prototype via son navigateur web, et ainsi envoyer des requêtes sur le serveur web.
- Le serveur web, qui reçoit les requêtes HTTP et transfère les requêtes reçues au serveur d'applications, transmet les résultats, qui sont ensuite affichées.
- Le serveur d'applications, qui contient les applications à exécuter et les données (ontologie). Il traite la requête en exécutant l'action choisie par le client et en manipulant les données stockées dans les fichiers, puis il envoie au serveur web le résultat.

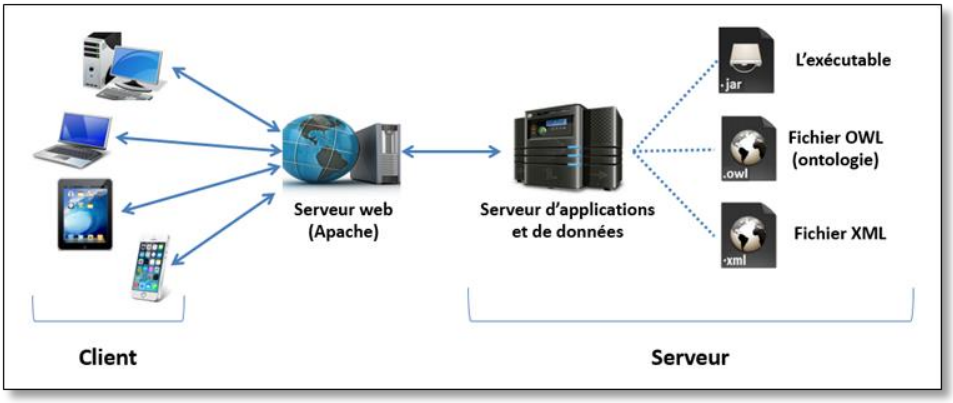


Figure 4-8. Architecture du module d'utilisation de la base de connaissances

4.1.3.1.2. La librairie Java OWL API

Afin de pouvoir gérer la lecture, l'écriture de concepts et d'instances dans une ontologie, ainsi que de contrôler le processus de raisonnement, la bibliothèque Java OWL API a été utilisée lors du développement de cette interface. Cette bibliothèque de code est open-source et propose une interface de programmation d'application (API) qui peut être utilisée afin de créer des classes et des méthodes, qui lisent, écrivent et enregistrent des fichiers OWL. Initiée en 2003, cette bibliothèque est aujourd'hui activement maintenue par l'université d'Ulm et par l'entreprise Clark & Parsia. Elle est encore utilisée dans de nombreux projets, tels que l'éditeur d'ontologies Protégé (dans sa version 4.3).

Cette API peut également être utilisée afin de traiter, modéliser et manipuler des données OWL, ainsi que pour effectuer une analyse fondée sur la logique (inférence qui analyse la sémantique des données contenues dans le fichier OWL). Pour cela, la bibliothèque OWL API est principalement utilisée pour développer des composants et des applications autonomes.

De plus, OWL API permet de mener un raisonnement à partir d'une ontologie, grâce à l'intégration de différents moteurs d'inférence, tels que Fact++, Pellet, Racer Pro, etc. Précédemment, lors de la présentation des règles SWRL et du processus de raisonnement, nous avons souligné notre choix pour le raisonneur Pellet, qui permet entre autres un raisonnement incrémental.

Les travaux de stage effectués au CRC par El Moussawi (El Moussawi 2014) nous ont permis de développer un module fondé sur Java, qui va rechercher l'ontologie OWL sur un serveur, la copier localement, la lire via une interface de consultation, et y intégrer de nouveaux objets. Les informations sur l'utilisation effectuée localement est ensuite envoyée, lors de la sélection des méthodes, via un protocole HTTP par des fichiers XML et le JDOM API (*Java Document Object Model*).

Les détails de cette solution d'échange d'informations sont expliqués dans le rapport de stage (El Moussawi 2014).

L'interface de ce module Java est présentée sur la Figure 4-9. On y retrouve les informations qui permettent de définir un objet *Situation* dans l'ontologie. Un premier cadre permet de saisir les informations relatives à l'utilisateur : un champ de texte pour décrire sa profession (à but de stockage et d'analyse) (zone 1), le choix d'un environnement de travail prédéfini dans l'ontologie (zone 2), ainsi que les informations sur son profil (barre de choix de la zone 3). Le choix de cette interface permet de simplifier le processus de modélisation d'un utilisateur, et ainsi tester plusieurs profils différents pour évaluer l'impact sur la sélection des visualisations. L'utilisation de cette interface est donc subjective, l'utilisateur jugeant par lui-même son niveau sur les différents axes proposés. Comme nous le verrons plus tard, cette modélisation du profil utilisateur peut être plus complexe.

Un second cadre permet de sélectionner les termes de l'ontologie qui décrivent l'analyse à mener. Cela correspond aux concepts *Risk*, *Goal* et *Task*, dont des instances sont définies au sein de l'ontologie. Par des menus déroulants, l'utilisateur peut ainsi choisir un type de risque à étudier (zone 4), ou bien préciser un but d'utilisation (zone 5) ou une tâche d'exploration précise (zone 6).

Comme nous l'avons décrit lors de la phase de modélisation, ces trois concepts successifs dépendent successivement du concept plus général. Ainsi, lors de la sélection d'un terme dans l'un de ces menus déroulants, les objets des niveaux suivants sont alors filtrés (voir la Figure 4-10). Dans la partie basse de cette interface, une place est laissée afin d'afficher les données nécessaires à ces opérations d'exploration et d'analyse. Celles-ci sont affichées lors de la sélection de l'opération à mener par la visualisation, et peuvent être sélectionnées par l'utilisateur selon la disponibilité de ces données.

Le bouton au bas de cette interface permet alors de valider la description du cas d'utilisation. Cette première étape de sélection correspond à la **lecture de l'ontologie**, tandis que la validation permet ensuite une **écriture dans l'ontologie**, enregistrée localement, sous la forme d'une nouvelle instance du concept *Situation*.

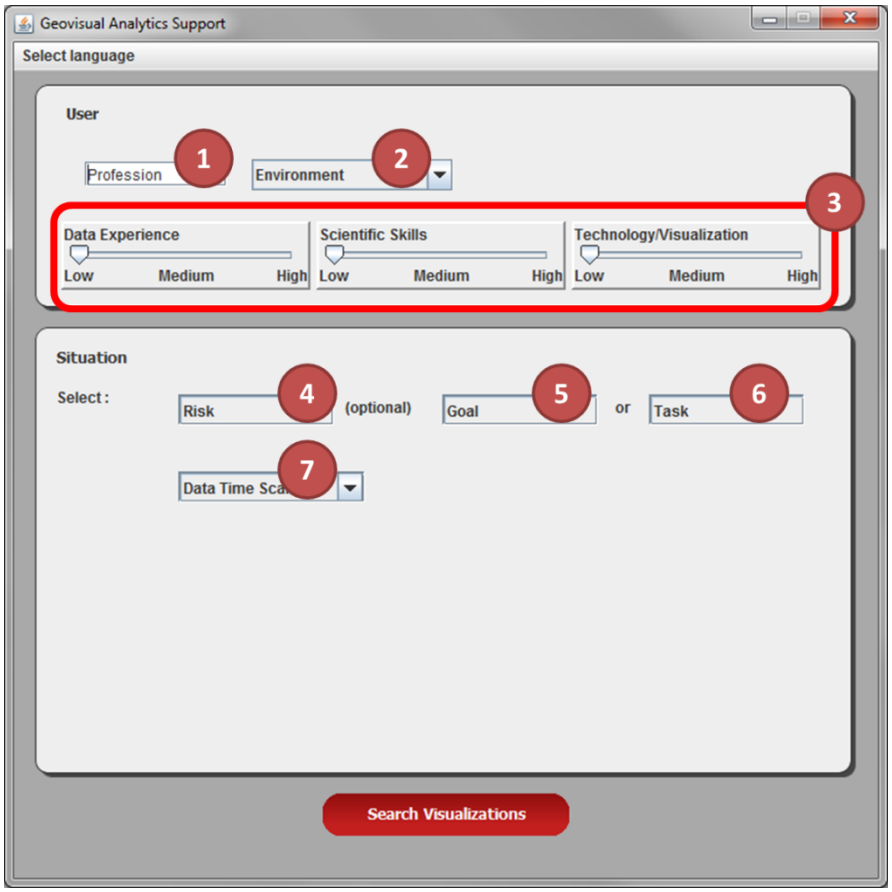


Figure 4-9. Interface Java pour l'utilisation de la base de connaissances

Figure 4-10. Sélection du but d'utilisation *Analyse de comportement de pêche,* puis affichage des différentes tâches et données correspondantes

Une fois la nouvelle instance *Situation* ajoutée dans l'ontologie locale, le processus de raisonnement est exécuté, grâce à ces nouvelles informations. Le résultat de cette démarche est alors une nouvelle fenêtre qui présente les méthodes de visualisation qui peuvent être utilisées, afin de définir un environnement complet d'analyse géovisuelle. Ces méthodes sont présentées sous plusieurs onglets méthodologiques, les trois espaces de Peuquet : géographique, temps et sémantique. Ce choix d'onglets a été fait afin de distinguer les méthodes selon cette triple taxonomie présentée dans les parties précédentes, et non pas selon une sémantique plus complexe qui aurait demandé une modélisation encore plus approfondie des méthodes de visualisation. Ces trois grandes classes de

visualisations n'étant pas exclusives, des méthodes peuvent se retrouver dans un ou plusieurs onglets à la fois, selon les données de base qu'elles utilisent.

Sur cette interface, seules les méthodes retenues sont représentées en couleur. Les autres méthodes, non sélectionnées par le système à base de connaissances, sont laissées en noir et blanc (voir Figure 4-11). Ces méthodes sont représentées par des icônes fixes, et non pas sur les données réellement étudiées. En passant la souris au-dessus de ces icônes, les images sont agrandies afin d'avoir une meilleure compréhension de la visualisation utilisée, et des détails relatifs à cette méthode sont données en commentaire (nom, description).

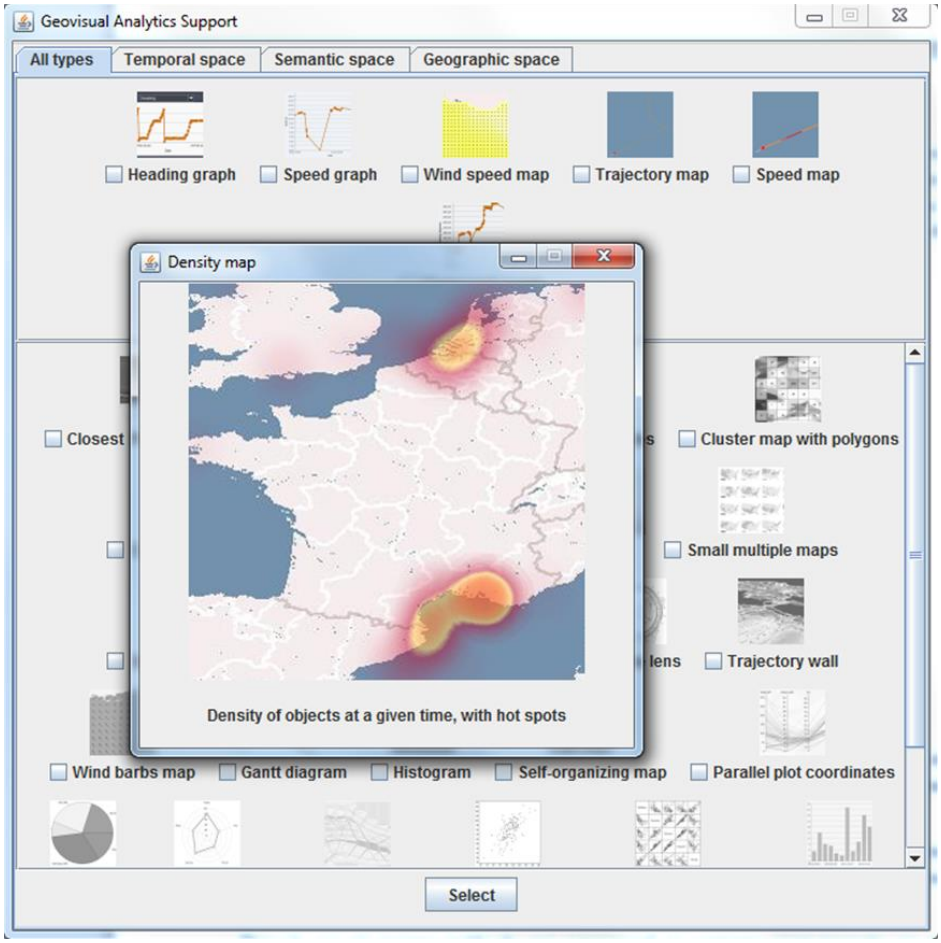


Figure 4-11. Fenêtre de choix des méthodes de visualisation après exécution du raisonneur dans l'ontologie

L'utilisateur peut alors choisir les méthodes de son choix à partir de cette interface, pour ensuite les utiliser sur les données réelles qu'il étudie. La plus grande partie des méthodes ainsi présentées sont intégrées dans la plateforme FishEye, par différents menus déroulants. Cette interface permet donc de conseiller quant à leur utilisation, selon l'analyse à mener. Un lien dynamique entre le choix de la méthode sur l'interface Java, et son application sur l'interface de FishEye, a pu être développé pour certaines de ces méthodes, comme preuve de concept. Toutefois, les méthodes ainsi modélisées sont parfois issues d'autres travaux de recherches, ou d'outils propriétaires, que nous n'avons pas pu intégrer à notre plateforme web. Parmi les 31 méthodes formalisées dans l'ontologie, 17 méthodes ont pu être intégrées sur cette plateforme d'analyse de données maritimes. La liste de ces méthodes de visualisations est présentées dans les annexes de ce manuscrit (voir l'annexe *2. Liste des*

visualisations formalisées, p.234). Celles-ci ont essentiellement été développées à partir des librairies JavaScript disponibles pour la visualisation d'information, libres et gratuites, telles que les librairies Ext, GeoExt ou D3. Les autres méthodes ont été testées indépendamment de l'application FishEye, car obéissant à des logiciels différents.

De par son langage de développement, cette interface Java peut alors être intégrée à un environnement web. La sous-partie suivante présente l'intégration de cette interface d'utilisation de l'environnement d'aide à l'analyse géovisuelle, dans la plateforme FishEye pour l'analyse des données maritimes.

4.1.3.2. Intégration du module à FishEye

Dans l'environnement FishEye, nous proposons trois manières d'accéder à cet outil d'aide à l'analyse géovisuelle. L'outil peut soit être chargé sans prendre en compte de données à étudier, depuis la barre d'outils (*Geovisual Analytics Support*). Toutefois, il est possible de sélectionner au préalable des données depuis la carte, afin que les méthodes choisies par la suite soient directement appliquées à ces données. Ainsi, nous permettons le choix d'une zone géographique ou d'un navire à analyser. La première option est possible en dessinant une zone (outil *Geovisual Analytics of Zone*, Figure 4-12 a), qui est alors considérée comme paramètre pour l'application des visualisations prochainement sélectionnées. La deuxième option est alors accessible par un clic droit sur un navire (Figure 4-12 b), dont on souhaite étudier le mouvement à l'aide des méthodes qui seront choisies par la suite.

(a) (b)

Figure 4-12. Accès à l'outil d'aide à l'analyse géovisuelle depuis FishEye : (a) par le menu général, (b) par un menu contextuel

Ce contrôle depuis l'interface de FishEye ouvre alors un lien vers le téléchargement de l'application Java. Celle-çi est alors copiée localement avec l'ontologie sur l'espace temporaire du client. Selon le choix de contrôle effectué, les informations relatives à la zone (étendue, coordonnées) ou au navire (coordonnées, types, etc.) sont gardées dans la mémoire de l'application web, qui les transfère à l'application Java. Une fois l'EAAG utilisé et les méthodes sélectionnées, celles-ci sont alors éxécutées dans l'interface de FishEye, afin d'étudier l'objet d'intérêt.

Afin d'intégrer une telle application à une page web et l'exécuter du côté client, Java dispose de deux technologies : les Applets et le Java Web Start. Une applet Java est un type particulier de programme Java qu'un navigateur compatible avec Java peut télécharger et exécuter. Une applet est typiquement incorporée à l'intérieur d'une page web et s'exécute dans le contexte d'un navigateur. *A contrario*, le Java Web Start permet de télécharger et d'exécuter des applications Java à partir du web. Le logiciel Java Web Start propose les avantages suivants :

- Il permet d'activer facilement les applications en un seul clic

- Il garantit l'exécution de la version la plus récente de l'application
- Il ne fait pas appel à des procédures d'installation ou de mise à niveau compliquées

Grâce au Java Web Start, les utilisateurs peuvent donc lancer une application Java en cliquant sur un lien dans une page web. Le lien pointe vers un fichier JNLP (*Java Network Launch Protocol*), qui indique à Java Web Start comment télécharger, cacher et exécuter l'application.

Les applets ne peuvent pas effectuer certaines opérations, telles que l'accès aux ressources des clients en mode écriture ; ce dont nous avons besoin pour cette application. Notre choix s'est donc porté sur la technologie Java Web Start qui permet de résoudre ce problème d'accès, tout en préservant la sécurité et l'intégrité du système (El Moussawi 2014). La Figure 4-13 montre le type de message que l'utilisation du JNLP fait alors apparaitre, avant de télécharger et exécuter l'application Java. Cette nouvelle fenêtre n'est donc pas une fenêtre du navigateur Internet, mais est une fenêtre indépendante. Grâce à JNLP, nous pouvons tout de même envoyer des informations d'une application à l'autre, via le serveur.

Dans cette partie, nous avons présenté l'interface homme-machine qui est fondamentale dans la définition de cet environnement d'aide à l'analyse géovisuelle : maintenant, nous présentons son utilisation et ses avantages. Dans la partie suivante, nous présentons donc un scénario d'intérêt pour l'exploration et l'analyse des données de mouvement dans le trafic maritime, pour lequel l'analyse géovisuelle et cet EAAG jouent un rôle fondamental.

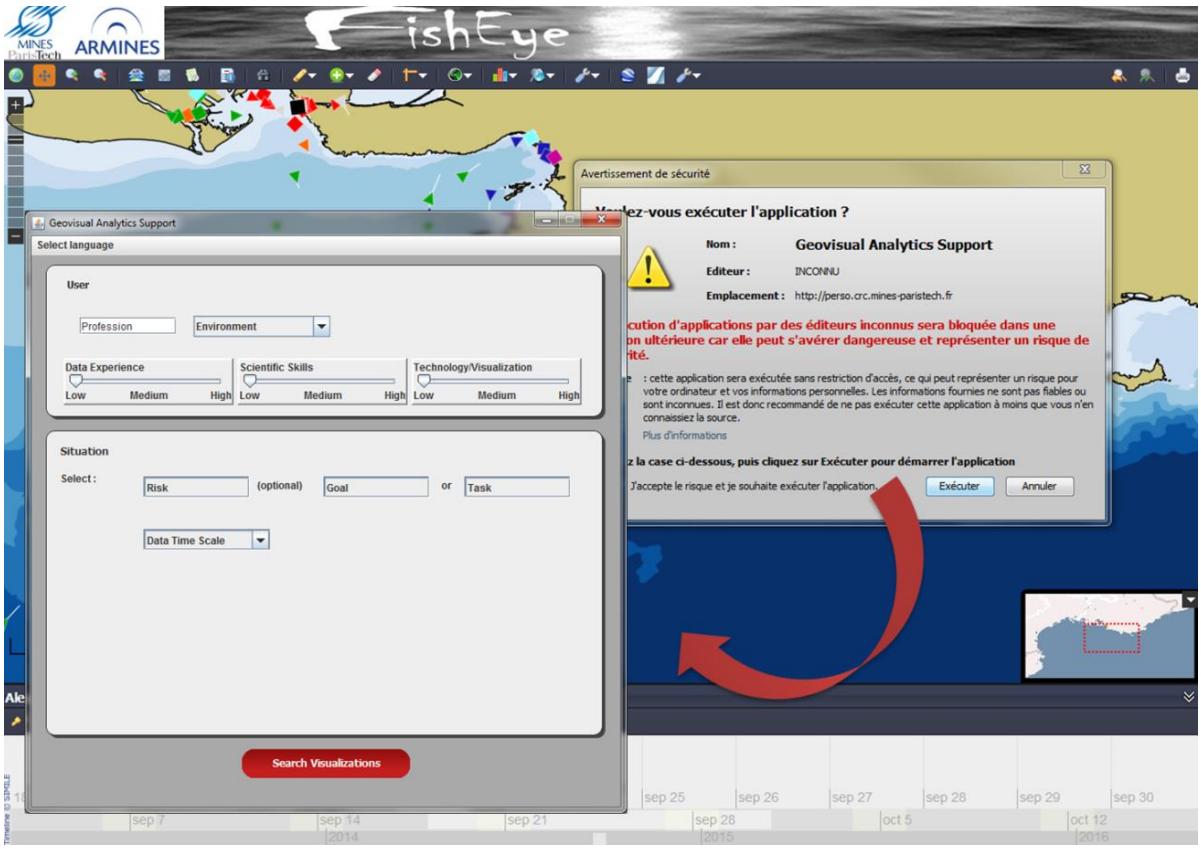


Figure 4-13. Exécution de l'aide à l'analyse géovisuelle depuis FishEye, et utilisation du JNLP

4.2. Expérimentation du module

4.2.1. Amélioration de la visualisation du trafic maritime

Nous pouvons rappeler les différents aspects jugés comme sensibles et nécessaires de visualiser, selon Davenport et Risley (Davenport et Risley 2006). Les auteurs mettent en avant les quatre groupes de tâches à améliorer par la visualisation, pour l'analyse et la surveillance intelligente du domaine maritime :

- Visualiser les informations qui nous sont **inconnues**, telles que les données manquantes et la qualité des informations déjà représentées
- Visualiser le comportement des navires et leurs déplacements à la fois **dans l'espace et dans le temps**, et intégrer de nouvelles manières (simples) de projeter l'information temporelle
- Visualiser les comportements dits **« normaux »** des navires, afin de mieux repérer les éléments anormaux, qui sortent de l'ordinaire
- Visualiser les **attributs d'intérêt** pour ce que les auteurs appellent le MISR (*Maritime Intelligence, Surveillance & Reconnaissance*)

Afin d'illustrer l'utilisation d'un environnement d'aide à l'analyse géovisuelle appliqué à ce domaine, nous choisissons donc d'explorer un scénario qui puisse porter sur un ou plusieurs de ces points fondamentaux. Comme nous l'avons vu au cours des chapitres précédents, les nombreuses recherches dans le domaine de l'analyse géovisuelle se sont majoritairement portées sur la visualisation et l'exploration de flux et de trajectoires, c'est-à-dire une analyse dans l'espace et dans le temps des mouvements d'objets. Ceci correspond notamment au deuxième point relevé par Davenport et Risley, tel que nous l'avons présenté ci-dessus.

Dans le cadre du domaine de la sécurité et sûreté de la maritimisation de l'énergie, nous choisissons d'étudier des situations de potentiels risques relatifs aux déplacements d'un tanker en mer. Un tanker, aussi connu comme navire-citerne, est un navire de charge destiné au transport de matières liquides, telles que le pétrole, les produits chimiques, ou encore du gaz liquéfié. L'étude de leurs mouvements et l'analyse de tout type de risques qui peuvent leur être associés, est alors fondamentale pour le contrôle des énergies sûres. Un accident qui implique un tanker peut en effet avoir des conséquences très graves sur l'environnement et sur l'économie d'une zone maritime ou d'un pays. Un comportement que l'on pourrait qualifier de « normal » pour un tanker en haute mer, est généralement caractérisé par une trajectoire rectiligne (cap constant) et une vitesse quasi-constante. La Figure 4-14 illustre par une carte une trajectoire habituelle de ce type de navire, dans la mer Méditerranée.

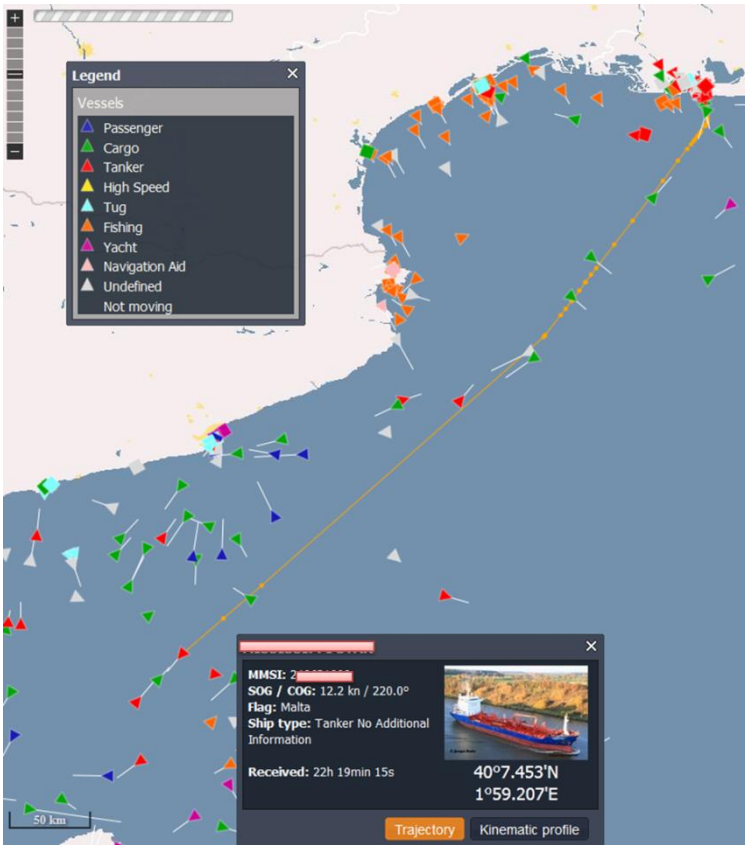


Figure 4-14. Trajectoire d'un tanker (navire rouge) au Sud-Est de l'Espagne, et fiche descriptive de son identité

Dans la sous-partie suivante, nous présentons un cas d'étude réel de collision de deux navires. Par la suite, nous présenterons plusieurs risques et anomalies possibles liés aux déplacements d'un tanker au sein de l'espace maritime, et proches de la situation présentée.

4.2.2. Un cas d'étude : collision de navires

En suivant le mantra de l'analyse visuelle, de telles méthodes doivent permettre de *« détecter l'attendu, et découvrir l'inconnu »* (Thomas et Cook 2005). Si l'on applique ce mantra à la gestion de risques, selon la classification Cynefin (Snowden et Boone 2007) des évènements (simple, compliqué, complexe, chaotique – voir Figure 4-15), nous pourrions changer cette phrase en : *l'analyse visuelle propose des méthodes pour reconnaitre des situations simples ou compliquées, et pour découvrir des relations dans des situation complexes*. Dans cette sous-partie, nous présentons spécialement un cas réel de collision, qui n'a pas pu être évité, mais qui a tout de même été identifié par le centre de surveillance relevant.

Nous présentons ici les différentes représentations cartographiques dans le cas d'une collision de navires, telles qu'elles ont pu être utilisées par les opérationnels et analystes. Par la suite, nous présentons des visualisations qui peuvent être proposées par notre approche d'analyse géovisuelle. Le cas étudié ici est une collision dans le pas de Calais, entre les navires Paula C (cargo) et Darya Gayatri (vraquier), le 11 Décembre 2013 (MAIB 2014). Le scénario est le suivant : le navire Paula C suivait une route en pilote automatique selon un cap de 221°, à une allure de 11,9 nœuds. Celui-ci était alors suivi par le navire Darya Gayatri, à un cap de 216° et une vitesse de 12,4 nœuds, menaçant de rattraper le Paula C (CPA nul sous 88 minutes). Les navires ont alors tenté une manœuvre

d'évitement en changeant leurs caps, mais un troisième navire (Raquel, chalutier) a alors été repéré sur leur nouvelle trajectoire. Suite à des problèmes de communication entre les trois navires, seul le Raquel a pu échapper à la collision, mais les deux navires Paula C et Darya Gayatri se sont rencontrés (198° et 12,9 nœuds pour le Darya, 70° et 6,2 nœuds pour le Paula C).

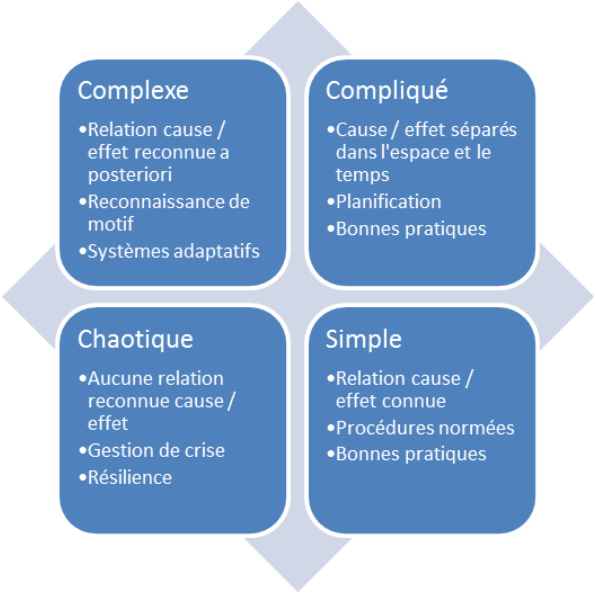


Figure 4-15. Classification Cynefin des évènements

Sur les Figure 4-16, Figure 4-17 et Figure 4-18, nous retrouvons les visualisations utilisées par les opérationnels pendant cet accident, ainsi quelle celles publiées après l'évènement pendant la phase d'analyse des faits. Ces images, issues du rapport du MAIB (*Marine Accident Investigation Branch*) représentent les outils de plusieurs acteurs au cœur de cet évènement, et montrent à la fois la diversité des représentations et les limites quant à l'interprétation des mouvements.

Ces images montrent que sans outil de projection, d'analyse ou de simulation, le contrôleur ne peut pas prédire trop en avance ce cas de collision. Bien souvent, de tels accidents sont repérés au cours des dernières minutes, au moment des manœuvres des navires pour empêcher la collision. Toutefois, le contrôleur en charge de la surveillance de ce trafic aurait pu identifier deux trajectoires proches et parallèles, un cas de figure intéressant car retrouvé au début de nombreux scénarios de risques. Une étude intéressante de ce cas d'accident serait alors de comprendre comment améliorer la visualisation côté contrôleur, de manière à ce qu'il puisse comprendre la situation, et dégager les risques possibles induits par ces trajectoires proches et parallèles.

Sur les représentations de l'information par des outils actuels pour la surveillance et pour la représentation du trafic maritime (Figure 4-16, Figure 4-17, Figure 4-18), la représentation de la dynamique des navires est absente, ou bien compliquée à interpréter. En effet, l'écran de contrôle (Figure 4-16) ne permet pas ici d'avoir une entière compréhension des trajectoires des navires, une seule position étant affichée par un vecteur vitesse. Toutefois, grâce à ces vecteurs, on remarque une possible zone de collision entre le Paula C, le Darya Gayatri et le Raquel. Sur la Figure 4-17, qui est une représentation de l'évènement a posteriori par un bureau d'enquête, afin d'illustrer le cas d'étude, la seule façon de comprendre la dynamique des navires est de supposer que leur émission AIS est régulière et identique, afin d'en déduire la vitesse relative par la succession de symboles navires (triangles). Toutefois, la notion de temps et d'espace est difficilement interprétable, car nous

ne pouvons pas savoir quelles sont les positions de chacun des navires, à un même instant t : les tâches de mesure de distance et de comparaison des vitesses est impossible par cette simple représentation des trajectoires.

Figure 4-16. Ecran de contrôle en centre de surveillance, les navires Paula-C et Darya Gayatri en navigation proche et parallèle, navire Raquel en manœuvre d'évitement

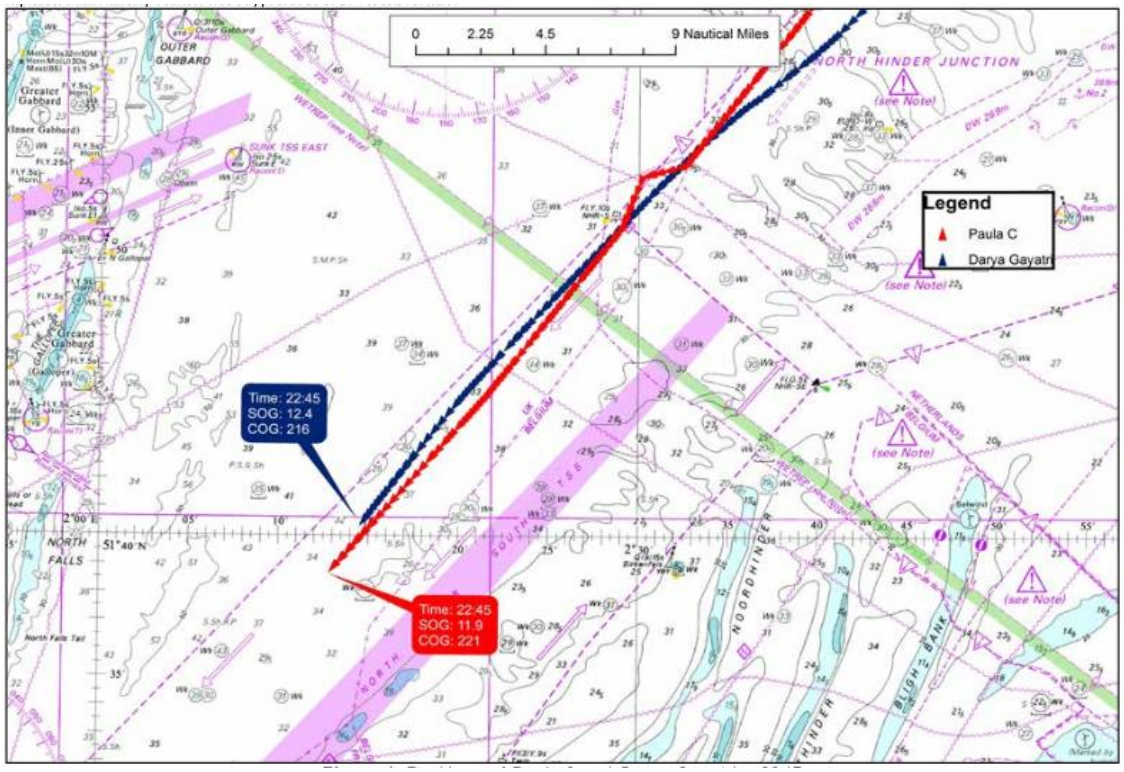


Figure 4-17. Trajectoires parallèles dans le pas de Calais, représentées après accident sur une carte électronique britannique

A titre d'exemple, la Figure 4-18 représente la vue radar à bord de l'un des navires au moment de l'accident. Cette représentation n'a pas fonction à comprendre la dynamique, mais montre la lourdeur de l'affichage des données sur cet écran. Dans le rapport d'accident, il est mentionné que cet outil radar était utilisé à bord du Paula C comme système anti collision, au lieu de l'utilisation habituelle des outils de l'ARPA (*Advanced Research Projects Agency*). De plus, les équipages ne savaient pas utiliser les données AIS, qui aurait pourtant permis d'identifier le troisième navire Raquel (pêche), dans la zone d'accident.

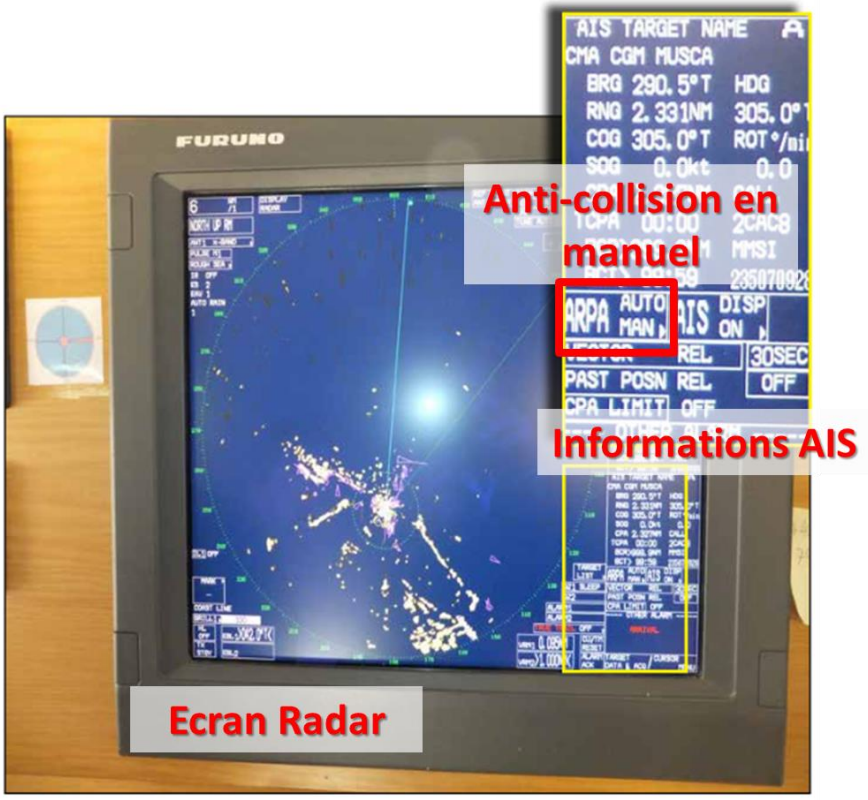


Figure 4-18. Ecran radar à bord du navire Paula C au moment de l'accident et affichage des données AIS

L'utilisation de ces outils destinés aux équipages des navires, par habitude les officiers de bord des deux navires, n'était donc pas adaptée à la situation dans laquelle ils se sont retrouvés, et donc ne présentait pas les bons outils pour visualiser les dangers proches. Sur l'écran radar utilisé, le minimum d'information sur l'identité des navires proches pouvait être affiché, et la dynamique des objets très compliquée à interpréter par ce seul écran.

En suivant les conseils de visualisation donnés par l'environnement d'aide à l'analyse géovisuelle dans une telle situation, du point de vue du contrôleur (test d'une hypothèse, visualisation des informations d'intérêt, confirmation) ou de l'analyste (analyse a posteriori de l'accident), nous représentons sur les figures suivantes l'utilisation de visualisations issues de notre environnement, dans le cas de trajectoires similaires à ce cas d'accident. La question se pose quant à l'interprétation de ce risque, selon la dynamique observée.

Grâce à la visualisation représentée sur la Figure 4-19, l'utilisateur peut utiliser des données en temps réel, qui mettent automatiquement à jour cette visualisation. L'utilisation de la carte pour représenter les trajectoires, juxtaposée à l'utilisation d'un graphe temporel qui compare les vitesses des navires, permet de valider le fait que ces deux navires ont une trajectoire en cours, proche, conservent la même vitesse moyenne de 10 nœuds. Aucun navire ne présente de variation de vitesse ou de cap important, comme pour une manœuvre d'évitement. Grâce à la cartographie, la distance entre ces deux navires peut être mesurée, comparée à des informations météorologiques, afin de conclure sur la possible visibilité des deux navires.

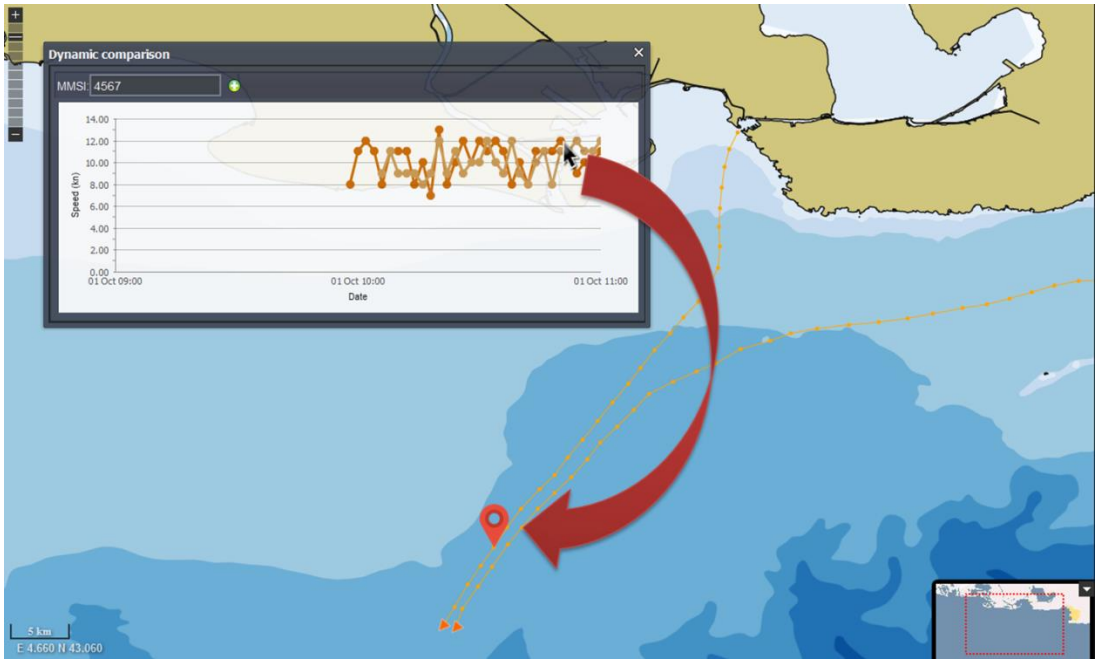
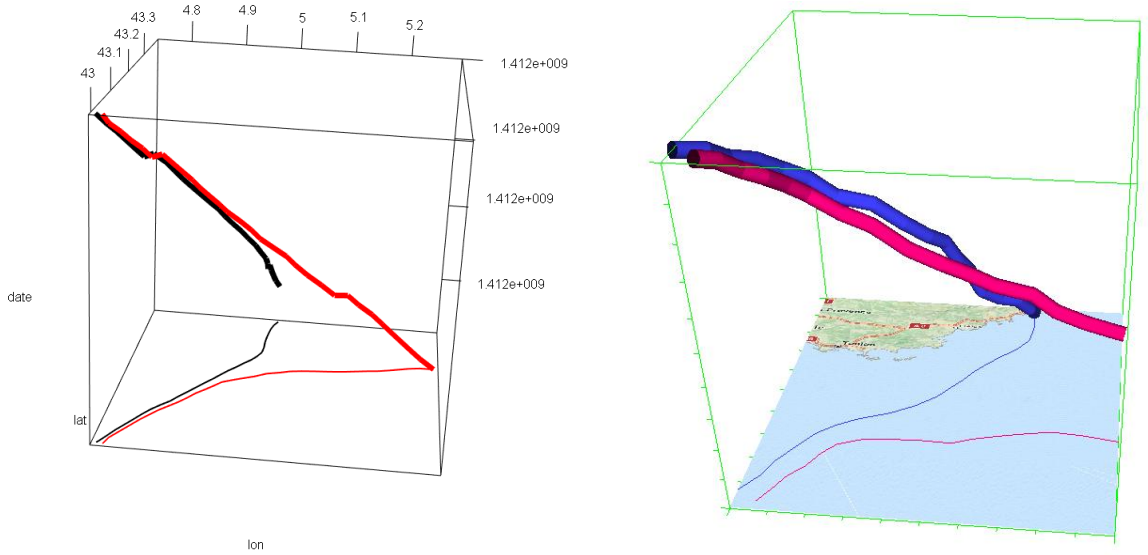


Figure 4-19. Représentation de trajectoires parallèles et graphe de vitesse correspondant

L'utilisation de cette représentation, en cas de données en temps réel, permet alors de vérifier les évènements suivants : les trajectoires sont linéaires et non en zigzag (pas d'attaque), la vitesse des deux navires est constante à 10 nœuds, les navires sont proches dans l'espace (<1km) et dans le temps (dernières positions enregistrées à 11h00 le même jour). Les deux trajectoires étant à cap constant, il ne s'agit pas d'un cas de dérive. Les situations de transbordement en cours ou de ravitaillement peuvent alors encore être testées, par un approfondissement de l'analyse.

Les représentations en trois dimensions sur la Figure 4-20 ne sont pas adaptées à du temps réel, à cause du temps de traitement de l'information, mais permettent d'analyser les mouvements passés de deux navires dans une unique visualisation. La pente des segments 3D permet d'estimer la vitesse relative des navires, et l'espace entre ces segments permet de savoir si les navires étaient proches dans l'espace et dans le navire, par une unique visualisation. Toutefois, les valeurs exactes de la vitesse ne sont pas accessibles sur cette représentation, et nécessite l'utilisation d'une visualisation supplémentaire, ou bien l'utilisation d'une variable visuelle. Ici, la couleur des *tubes* permet de distinguer les deux navires (selon leur identifiant MMSI), mais nous pourrions penser à utiliser cette couleur ou la largeur des tubes, afin de représenter la vitesse.



(a) (b)

Figure 4-20. Représentation de trajectoires de navires parallèles par un cube spatio-temporel grâce (a) au logiciel R et (b) au logiciel ILWIS

4.2.3. Exploration et analyse des mouvements de navires

De manière à illustrer l'utilisation de l'EAAG pour des cas plus importants, nous proposons plusieurs cas d'utilisations où les scénarios à étudier sont dits complexes, selon la classification Cynefin : c'est-à-dire que seule une analyse *a posteriori* va permettre d'identifier, et comparer l'évènement enregistré à des motifs connus. Par définition de cette classe, un tel type d'évènement complexe ne peut donc pas être reconnu en temps réel par les systèmes de surveillance maritime actuel. En suivant l'exemple présenté précédemment, nous présentons ici les différentes interprétations qui peuvent être faites de deux navires ayant deux trajectoires proches et parallèles.

4.2.3.1. Présentation des scénarios complexes

Sur la Figure 4-21, nous présentons les formes de trajectoires qui peuvent correspondre à quatre scénarios différents dans le cas de trajectoires proches et parallèles. Les simples formes des trajectoires dans ces quatre cas sont très ressemblantes, et il est compliqué de donner du sens aux mouvements de ces objets. La différence de comportements sur ces cas ne peut être jugée que par une analyse approfondie des informations relatives à l'identité et à la dynamique du mouvement. Les différents cas de situations présentés ici sont :

- **Scénario 1 :** le tanker A est attaqué par le navire pirate B
- **Scénario 2 :** le navire A est en dérive, le navire B l'approche pour le secourir
- **Scénario 3 :** deux tankers A et B effectuent un transbordement
- **Scénario 4 :** le navire A est suivi par B, les deux risquent de rentrer en collision

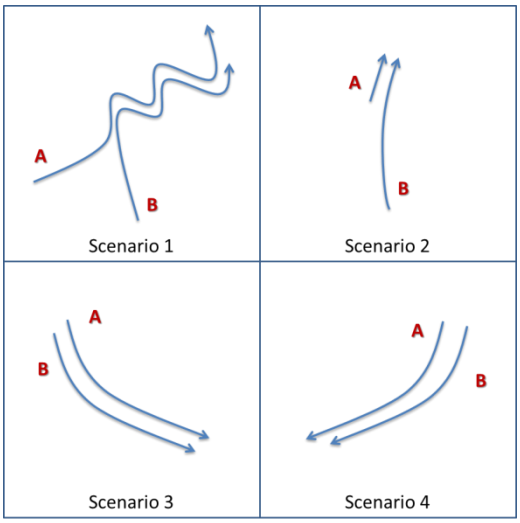


Figure 4-21. Formes de trajectoires suspectes lors de l'approche d'un navire

Ces différents scénarios peuvent alors être distingués par l'analyse du mouvement dans l'espace et le temps, dans le contexte géographique et météorologique, ainsi que par l'identification des navires étudiés (leur type, destination, etc.). Dans la sous-partie suivante, nous présentons en détail chacun de ces scénarios, de manière à relever les informations nécessaires à prendre en compte dans l'analyse des données, et reconnaitre ces cas de risque

4.2.3.1.1. Attaque de A par B

Le tanker A avance à une allure moyenne de 10 nœuds, selon une trajectoire linéaire (cap constant), telle qu'illustrée sur la Figure 4-14 pour un tanker. Un petit navire B, qui n'envoie pas de données AIS, se rapproche à allure rapide (~20 nœuds) du navire A. De manière à éviter l'abordage, le navire A adopte une trajectoire en « zigzags » (changements de cap importants). Le navire B tente de se rapprocher et adopte la même trajectoire, en conservant une allure similaire au navire B. Si l'abordage est réussi, le navire est généralement stoppé en pleine mer. Ce scénario doit prendre place dans une zone peu fréquentée, où d'autres navires ne doivent pas voir l'action d'abordage. De plus, pour permettre une telle attaque, la mer doit être calme.

4.2.3.1.2. Dérive de A, aide de B

Le tanker A est en dérive, à une vitesse faible qui est proche de celle des courants. Son cap n'est pas fixe, et suit la direction des courants. Le navire ne diffuse plus de messages au centre de contrôle. Le navire B repère cette dérive, et vient se rapprocher du tanker A : il adopte une vitesse proche de ce navire, et une direction tangente aux mouvements. L'état de la mer et la fréquentation de la zone n'importe pas sur ce scénario, une dérive pouvant arriver pour de nombreuses raisons, et dans n'importe quel contexte.

4.2.3.1.3. Transbordement entre A et B

Les deux navires A et B sont des tankers, qui transportent des marchandises ou du pétrole. Ils proviennent de deux endroits différents, et adoptent des trajectoires similaires. Une fois proches et visibles l'un de l'autre, ils conservent une allure normale (~10 nœuds) et un cap constant. Les deux navires doivent être proches pour permettre l'échange de marchandises (inférieur à une distance de manœuvre D_T) et naviguer de cette façon pendant un laps de temps défini (T_T ~1h). Après cette

période d'échange, les deux navires se séparent (changement de cap). De plus, la mer doit être calme afin de permettre un transbordement.

Un transbordement peut être un acte licite ou illicite. Dans le second cas, il peut s'agir de l'équipage d'un tanker qui en prend contrôle et fournit la cargaison vers un autre tanker. Autrement, un tel acte illicite de transbordement peut être précédé d'une attaque, telle que décrite dans la partie 4.2.3.1.1, suivie de l'arrivée d'un second tanker « pirate ».

4.2.3.1.4. Collision entre A et B

Dans ce cas de risque, la météorologie et le contexte de navigation n'ont pas d'influence directe sur la collision. Toutefois, une information telle que le manque de visibilité entre les deux navires pourrait expliquer un tel accident. Les deux navires ont des trajectoires linéaires et identiques. Le navire B se rapproche progressivement du navire A (vitesse plus élevée), comme nous l'avons décrit dans le cas réel d'accident dans la partie 4.2.2 de ce chapitre. La collision peut avoir peut après un unique changement de cap d'au moins un des deux navires, dans des directions opposées, lors d'une manœuvre d'évitement.

Selon ces quatre différents scenarios, la reconnaissance ou non d'une menace s'appuie donc essentiellement sur le type des navires étudiés, le contexte météorologique et la dynamique passée des trajectoires, telle que la vitesse et les changements de cap successifs des navires. Dans la sous-partie suivante, nous présentons les modélisations effectuées pour la stratégie de d'analyse visuelle à adopter.

4.2.3.2. Formalisation des scénarios et du cas d'utilisation

Dans le cadre de notre formalisation du processus d'analyse géovisuelle, et des solutions qui permettent une telle exploration, les scénarios présentés précédemment peuvent être modélisés selon les évènements qui doivent être reconnus dans les données par la visualisation, afin que l'opérateur en charge de cette exploration identifie le possible scénario en cours. Au sein de l'ontologie, il s'agit des relations sémantiques entre les concepts *Risk*, *Goal* et *Task*. Les scénarios présentés sont des instances du concept *Risk*, qui est alors identifié par une succession de tâches d'analyse. Nous présentons ici la modélisation des processus d'analyse géovisuelle qui permette de reconstruire les scénarios de risques précédents, grâce aux évènements qui permettent de les identifier. Ceux-ci sont listés sur la Table 4-1, ainsi que des mots clés relatifs aux tâches d'analyse géovisuelle qui permettent de reconnaitre ces évènements. Sur cette table, il s'agit des évènements et tâches de visualisation qui peuvent être reconnu après un certain temps, particulièrement après l'établissement des trajectoires parallèles. De manière à compléter cette approche, la Table 4-2 reprend le cas du scénario d'attaque selon une approche en temps réel, en intégrant les tâches dans l'ordre d'identification des évènements.

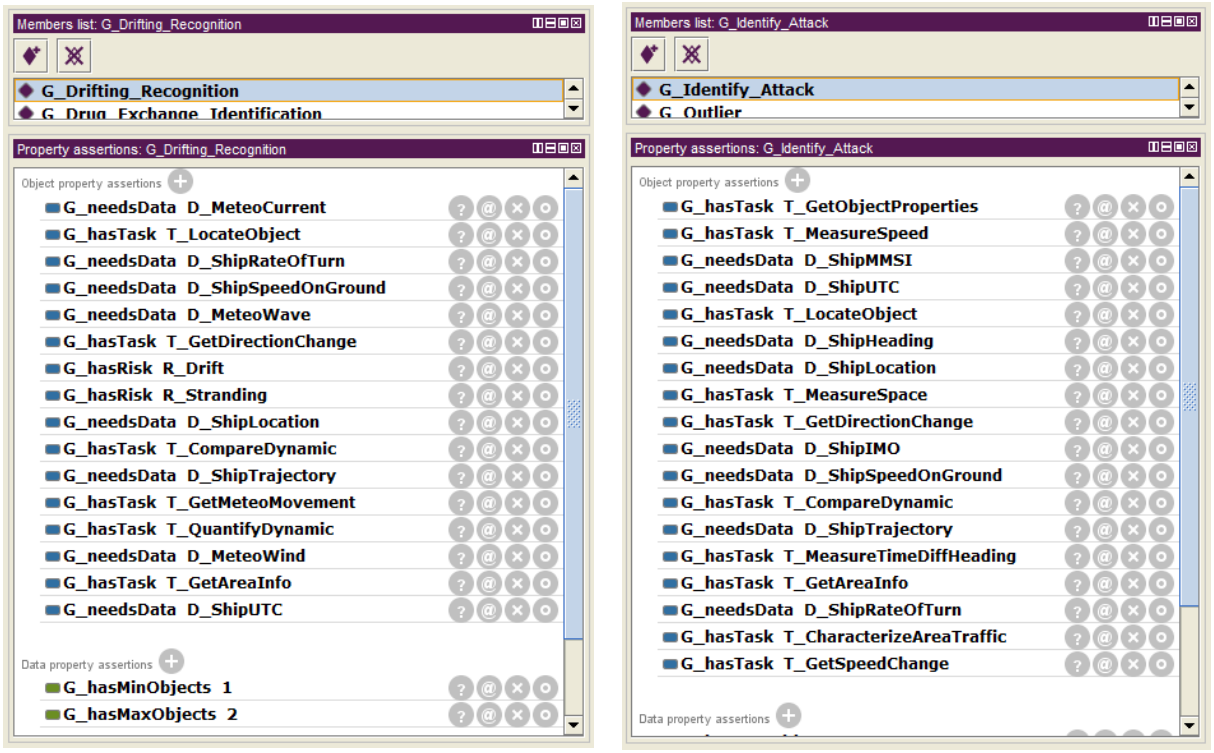
Table 4-1. Différenciation des scénarios identifiés, par la visualisation d'information a posteriori

Scénario	Evènements et données	Tâches GeoVA
Scénario 1 Attaque A par B	**A :** AIS, **B :** pas d'AIS	Identifier la source de données
	A&B : trajectoire en zigzag	Visualiser la forme de la trajectoire
	A : vitesse 10 nœuds	Visualiser l'évolution de la vitesse
	B : vitesse rapide, puis même que A	Comparer les vitesses
	Mer calme	Récupérer les informations de l'état de la mer
	Zone peu fréquentée	Caractériser l'état du trafic
Scénario 2 Dérive A	**A :** cap selon les courants	Visualiser la direction des courants
	A : vitesse faible selon les courants	Visualiser l'évolution de la vitesse
		Comparer la vitesse aux courants
	B : vitesse élevée, puis même que A	Comparer la vitesse
	B : cap moyen de A	Visualiser la forme de la trajectoire
		Comparer les caps suivis
	Distance faible	Mesurer la distance
Scénario 3 Transbordement A vers B	**A&B :** type tanker	Identifier le navire
	A&B : origines différentes	Récupérer les informations de route Visualiser la forme de la trajectoire
	B : approche de A, puis même vitesse que A	Visualiser la forme de la trajectoire
		Comparer les vitesses
		Localiser un lieu de rencontre
	Distance faible pendant un temps long	Mesurer la distance
	Contexte : Mer calme	Récupérer les informations de l'état de la mer (contexte)
Scénario 4 Collision A par B	**A&B :** même cap	Comparer les caps suivis
		Visualiser la forme de la vitesse
	B plus rapide que A	Comparer les vitesses
	B : approche de A	Mesurer la distance
		Localiser un lieu de rencontre
	A&B : au moins un évitement	Comparer les caps suivis
	A&B : vitesse faible après rencontre	Comparer les vitesses

Table 4-2. Ordre d'identification des évènements pour le scénario 1, Attaque de pirates

Phases chronologiques	Evènements et données	Tâches GeoVA
Phase 1 : Rien d'anormal	**A :** AIS, **B :** pas d'AIS	Identifier la source de données
Phase 2 : Approche	**A :** trajectoire rectiligne et vitesse 10 nœuds	Visualiser la forme de la trajectoire Visualiser l'évolution de la vitesse
	B : trajectoire rectiligne et vitesse rapide	Visualiser la forme de la trajectoire Visualiser l'évolution de la vitesse
	A&B : approche en cours	Visualiser la forme de la trajectoire Mesurer la distance Identifier une zone de rencontre
Phase 3 : Manœuvres	**A&B :** trajectoire en zigzag	Visualiser la forme de la trajectoire
	A devance B	Mesurer une différence de temps
	B : vitesse similaire à A	Comparer les vitesses
Phase 4 : Abordage	**A&B :** vitesses quasi-nulles	Comparer les vitesses Repérer un arrêt
	A&B : accolés	Mesurer la distance
	Contexte : Mer calme Peu de navires	Récupérer les informations de l'état de la mer Caractériser l'état du trafic

Au sein de l'ontologie, nous modélisons donc les processus d'analyse qui permettent de reconstituer ces différents scénarios, selon l'interprétation que l'utilisateur peut faire des données identifiées. Cette modélisation n'est pas celle du scénario de risque, mais des étapes qui permettent de reconnaitre les évènements propres à chaque scénario. La Figure 4-22 représente la modélisation ontologique des instances de *Goal* relatives à la reconnaissance de ces différents cas, par l'utilisation de différentes tâches d'analyse géovisuelle.



(a) (b)

Figure 4-22. Modélisation des processus d'analyse pour les cas (a) de dérive et (b) d'attaque pirate

Afin d'illustrer l'influence du changement d'utilisateur, nous prenons comme exemples les deux cas d'utilisation suivants, pour des utilisateurs professionnels du monde maritime, dans deux environnements distincts :

- Un contrôleur identifie deux navires proches, mais ne sait pas s'il y a une situation dangereuse ou non. Les deux navires peuvent être en situation irrégulière ou non, il doit donc effectuer un contrôle de ces données en temps réel pour mieux comprendre la situation. Ce contrôleur travaille dans un centre de surveillance maritime, dont l'environnement et le temps nécessaire requièrent des visualisations simples et rapides à obtenir. Dans le cas où cet utilisateur manipule l'EAAG, des hypothèses sur le risque à évaluer doivent être effectuées, afin de rechercher des méthodes de visualisation qui permettent de reconstituer ce risque.
- Un analyste doit étudier des trajectoires passées, où deux trajectoires de navires ont été identifiées comme proches et parallèle pendant une heure. Les navires ont pu avoir un comportement illicite ou non. L'étude des données doit permettre de comprendre si un tel acte illicite a effectivement eu lieu. L'analyste travaille en bureau, et peut avoir accès à des processus d'analyse plus complexe, sur des données à plus long terme.

Chacun de ces cas d'utilisation correspond alors à un objet *Situation* modélisé dans l'ontologie. La Figure 4-23 représente le premier cas d'utilisation, où un utilisateur de type *contrôleur* effectue une hypothèse sur un risque à évaluer, et avoir les visualisations qui lui permettent d'infirmer ou non ce scénario. Cette figure présente le cas lié à un scénario de dérive.

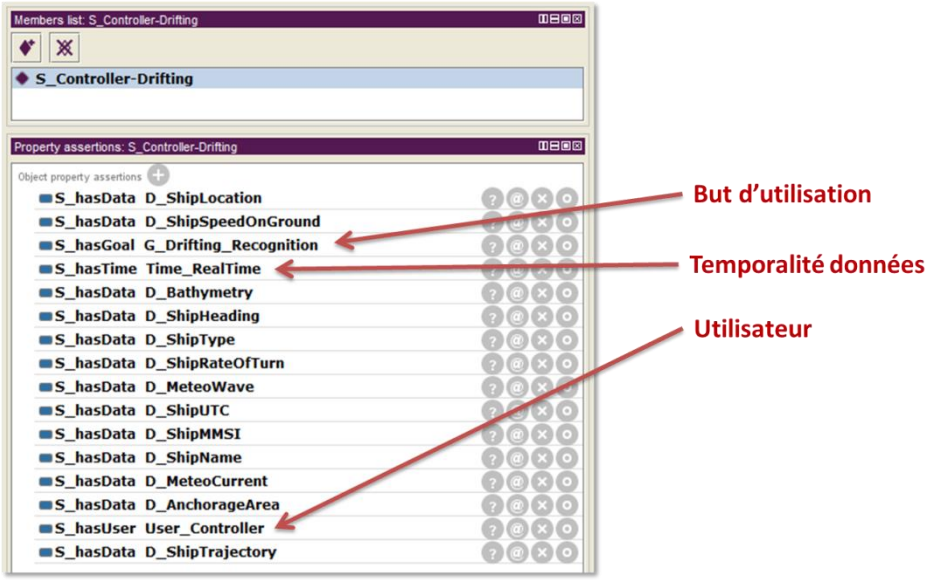


Figure 4-23. Cas d'un contrôleur amené à tester une hypothèse de dérivé : prise en compte de son profil, du but, de la temporalité des données et des données disponibles

4.2.3.3. Résultats et visualisations

Suite à la formalisation de ces cas d'utilisation, le raisonneur est utilisé afin de rechercher des visualisations qui puissent satisfaire ces critères de tâches d'analyse, de données, de temporalité de l'information et de profil d'utilisateur. Pour cela, l'ensemble des règles présentées dans la partie 2.3.3.1 sont appliquées successivement afin de créer de nouvelles dépendances entre les objets, lorsque les conditions des règles sont observées. Le raisonneur choisi permet de boucler l'application de ces règles de façon intérative, jusqu'à ce qu'un état d'équilibre soit observé ; c'est-à-dire qu'aucune contradiction n'est observée suite à l'application des relations nouvellement créées (règles contradictoires).

Les résultats, présentés depuis l'éditeur de l'ontologie en jaune, sont les axiomes ajoutés à l'objet de *Situation* pour représenter les visualisations adaptées au cas spécifique. Sur la Figure 4-24, on présente les résultats dans le cas d'un contrôleur qui souhaite évaluer un possible scénario de dérive.

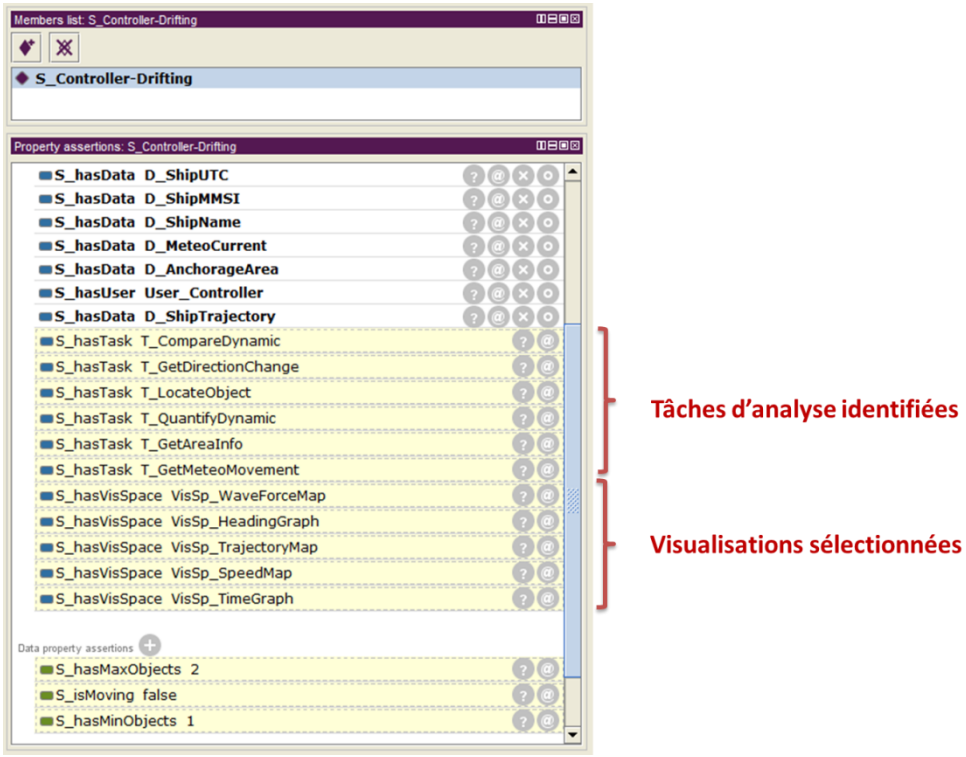


Figure 4-24. Résultats du raisonnement pour le cas d'un contrôleur et d'un scénario de dérive

Afin de reprendre les résultats du raisonneur liés à chacun des cas identifiés ici, la Table 4-3. Proposition de visualisations selon les utilisateurs et les scénarios étudiés liste les différentes méthodes de visualisation proposées selon l'utilisateur et selon le scénario testé. Par la suite, nous comparons alors les avantages de ces méthodes de visualisation proposées, par rapport à un système de surveillance classique.

Table 4-3. Proposition de visualisations selon les utilisateurs et les scénarios étudiés

Scénario / Utilisateur	Scénario 1 Attaque	Scénario 2 Dérive	Scénario 3 Transbordement	Scénario 4 Collision
Contrôleur (temps réel)	*WaveForceMap* *HeadingGraph* *TrajectoryMap* *SpeedMap* *StopMap* *DensityMap*	*WaveForceMap* *HeadingGraph* *TrajectoryMap* *SpeedMap*	*HeadingGraph* *TrajectoryMap* *CPA* *SpeedMap*	*HeadingGraph* *TrajectoryMap* *CPA* *SpeedMap*
Analyste (a posteriori)	- (temps reel)	*WaveForceMap* *HeadingGraph* *TrajectoryMap* *SpeedMap* *Gantt* *DriftSimulation*	*HeadingGraph* *FlowMap* *TrajectoryMap* *CPA* *ParallelCoordinates* *SpaceTimeCube* *SpeedMap* *Gantt*	*HeadingGraph* *FlowMap* *TrajectoryMap* *CPA* *SpaceTimeCube* *SpeedMap* *Gantt*

Si l'on reprend les évènements listés précédemment, afin d'identifier les différents scénarios, nous pouvons reconnaitre des éléments facilement reconnaissables par l'analyse géovisuelle, tandis que l'utilisation de systèmes d'identification et de surveillance actuels (ECDIS, données AIS, écrans radar, etc.) les rend plus difficiles d'accès (outils de mesure, de consultation, d'analyse) et compliqués à représenter. La Table 4-4 liste ces possibilités, pour trois des quatre scénarios identifiés (les scénarios 3 et 4 étant quasiment identiques), pour les utilisateurs contrôleurs et analystes, et selon les outils identifiés dans le cas précédent. Par l'utilisation des règles et de la modélisation développée, l'ensemble des méthodes de visualisation retournées par l'environnement d'aide à l'analyse géovisuelle permettent, par défaut, toutes les tâches demandées.

Table 4-4. Evènements reconnus par certains systèmes actuels étudiés et par la GeoVA

Scénario	Tâche GeoVA	Systèmes actuels		GeoVA
		Contrôleur	Analyste	
Scénario 1 Attaque A par B	Identifier la source de données	✓	✓	✓
	Visualiser la forme de la trajectoire		✓	✓
	Visualiser l'évolution de la vitesse			✓
	Comparer les vitesses			✓
	Caractériser l'état du trafic		✓	✓
	Récupérer les informations sur l'état de la mer	✓	✓	✓
Scénario 2 Dérive A	Visualiser la direction des courants	✓	✓	✓
	Visualiser l'évolution de la vitesse			✓
	Comparer la vitesse aux courants			✓
	Comparer la vitesse			✓
	Visualiser la forme de la trajectoire		✓	✓
	Comparer les caps suivis		✓	✓
	Mesurer la distance	✓	✓	✓
Scénario 3 Transbordement A vers B	Identifier le navire	✓	✓	✓
	Récupérer les informations de route	✓	✓	✓
	Visualiser la forme de la trajectoire		✓	✓
	Comparer les vitesses			✓
	Localiser un lieu de rencontre	✓	✓	✓
	Mesurer la distance	✓	✓	✓
	Récupérer les informations de l'état de la mer	✓	✓	✓
Scénario 4 Collision A par B	Comparer les caps suivis			✓
	Visualiser la forme de la vitesse			✓
	Comparer les vitesses			✓
	Mesurer la distance	✓	✓	✓
	Localiser un lieu de rencontre	✓	✓	✓
	Comparer les caps suivis			✓
	Comparer les vitesses			✓

Toutefois, nous n'avons pas ici conscience de la temporalité des phases successives dans la reconnaissance de ces scénarios. Si l'on reprend le cas du scénario 1, au cours du temps et des différentes phases identifiées sur la Table 4-2, nous pouvons ainsi identifier les différentes visualisations proposées étape par étape, et les phases qui peuvent donc être reconnues ou non au cours du temps, dans le cas de la surveillance en temps réel de ces informations. La Table 4-5 approfondit ce cas de scénario d'attaque, dans le cas d'un contrôleur qui surveille le trafic maritime

en temps réel. Sur cette table, l'interprétation de chaque phase doit se faire grâce à l'identification des phases précédentes : en effet, la simple identification de deux navires accolés et à vitesses quasi-nulles (phase 4) n'est pas nécessairement le signe d'un abordage, si les phases d'approche et de manœuvres n'ont pas été reconnues.

Dans le cas du contrôleur qui surveille le trafic maritime en temps réel, grâce à aux systèmes actuels, on peut relever les limites sur les phases suivantes, si celles-ci sont **considérées de façon indépendantes** (i.e., sans avoir repéré la phase précédente) :

- Phase 1 : la comparaison des outils AIS et Radar peuvent permettre d'identifier les différences de provenance des navires affichés.
- Phase 2 : difficulté de repérer la forme des trajectoires et identifier un cas d'approche, cela de manière rapide. Deux navires doivent être sélectionné, si repérés, avant de pouvoir calculer leur CPA.
- Phase 3 : difficulté à comparer les vitesses de deux navires et avoir une vue globale de leurs mouvements, dans l'intégralité des mouvements des navires surveillés.
- Phase 4 : si l'arrêt n'est pas mis en évidence, seules les positions proches de deux navires peuvent être repérées sur l'interface de surveillance.

Table 4-5. Scénario d'attaque, surveillé par un contrôleur maritime, et étapes de visualisation dans le temps

Phases chronologiques	Evènements et données	Systèmes actuels	Méthodes proposées
Phase 1 Rien d'anormal	**A :** AIS **B :** pas d'AIS Trafic régulier	✓	(pas de tâche d'analyse temps réel)
Phase 2 Approche	**A :** vitesse 10 nœuds		*HeadingGraph* *TrajectoryMap* *SpeedMap* *SpeedGraph*
	A : vitesse 10 nœuds	✓	
	B : trajectoire rectiligne		*HeadingGraph* *TrajectoryMap* *SpeedMap* *SpeedGraph*
	B : vitesse rapide	✓	
	A&B : approche en cours		*TrajectoryMap* *CPA*
Phase 3 Manœuvres	**A&B :** trajectoire zigzag		*HeadingGraph* *TrajectoryMap*
	A devance B	✓	*SpeedGraph* *TrajectoryMap*
	B : vitesse similaire à A		*SpeedGraph* *SpeedMap*
Phase 4 Abordage	**A&B :** vitesses quasi-nulles		*SpeedGraph* *StopMap*
	A&B : accolés	✓	*TrajectoryMap*
	Contexte : Mer calme	✓	*WaveForceMap* *DensityMap*
	Zone peu fréquentée	–	

Grâce à la Table 4-5, nous pouvons reconnaitre le nombre important de méthodes d'analyse géovisuelle qui peuvent être utilisées à chaque étape de ce raisonnement personnel, lors de l'analyse d'informations de mouvement. De par le profil de contrôleur et **l'étude des données en temps réel**, qui contraint l'utilisation de visualisations avancées, les méthodes proposées sont essentiellement les plus faciles d'interprétation et les plus rapides à utiliser dans un contexte de gestion d'alertes. Toutefois, il est intéressant de comparer à l'approche que pourrait avoir l'analyste pour l'analyse de ces données passées, et l'étude des différentes phases de ce scénario.

Table 4-6. Scénario d'attaque, étudié par un analyste

Phases chronologiques	Evènements et données	Méthodes proposées
Phase 1 Rien d'anormal	**A :** AIS **B :** pas d'AIS Trafic régulier	*ParallelCoordinates* *KohonenMap* *DensityMap* *CompositeDensityMap* *TrajectoryWall*
Phase 2 Approche	**A :** vitesse 10 nœuds	*HeadingGraph* *TrajectoryMap* *SpeedMap* *SpeedGraph* *CPA* *SpaceTimeCube* *FlowMap*
	A : vitesse 10 nœuds	
	B : trajectoire rectiligne	
	B : vitesse rapide	
	A&B : approche en cours	
Phase 3 Manœuvres	**A&B :** trajectoire zigzag	*HeadingGraph* *TrajectoryMap* *SpeedGraph* *SpeedMap*
	A devance B	
	B : vitesse similaire à A	
Phase 4 Abordage	**A&B :** vitesses quasi-nulles	*SpeedGraph* *StopMap* *TrajectoryMap* *WaveForceMap* *DensityMap* *CompositeDensityMap* *SpaceTimeCube*
	A&B : accolés	
	Contexte : Mer calme	
	Zone peu fréquentée	

De la même manière, la Grâce à la Table 4-5, nous pouvons reconnaitre le nombre important de méthodes d'analyse géovisuelle qui peuvent être utilisées à chaque étape de ce raisonnement personnel, lors de l'analyse d'informations de mouvement. De par le profil de contrôleur et **l'étude des données en temps réel**, qui contraint l'utilisation de visualisations avancées, les méthodes proposées sont essentiellement les plus faciles d'interprétation et les plus rapides à utiliser dans un contexte de gestion d'alertes. Toutefois, il est intéressant de comparer à l'approche que pourrait avoir l'analyste pour l'analyse de ces données passées, et l'étude des différentes phases de ce scénario.

Table 4-6 présente les méthodes proposées pour le même scénario, mais pour un utilisateur plus averti dans le domaine de la visualisation d'information, et pour des données passées (rôle de l'analyse). Les systèmes possibles sont nombreux et dépendent souvent du bureau d'étude chargé de l'analyse de ces données, comme nous avons pu en présenter certains dans les parties précédentes. Contrairement au premier tableau, les méthodes proposées par l'analyse géovisuelle dans ce cas permettent de prendre en compte l'intégralité du trafic maritime passé, avant et après l'évènement, de manière à mieux comprendre le contexte des mouvements.

Mais nous pouvons aussi identifier certaines limites d'une démarche analyse géovisuelle, limitée aux méthodes que nous avons pu formaliser. Par exemple, lors de la phase 2 (identifier une approche en cours), certains des outils proposés doivent être utilisés une fois que le contrôleur **a identifié des navires suspects**, et cherche à savoir si ces objets peuvent effectivement se rencontrer ou non. Il s'agit avant tout d'un outil de simulation (prévoir le futur), et non d'analyse à proprement parler (comprendre le passé et son contexte). Il est donc intéressant de repérer quelles étapes le contrôleur peut valider ou non, et s'il peut remonter à un scénario d'attaque ou non, grâce aux systèmes actuels ou par l'analyse géovisuelle.

- Si seule la phase 4 est repérée, par l'identification de deux navires immobilisés dans une zone où peu de navires sont présents, l'opérateur n'a pas accès à la dynamique des manœuvres qui ont précédé cet évènement. Il ne peut donc pas faire d'hypothèse sur une possible attaque ou non. Toutefois, l'utilisation de visualisations supplémentaires lui permettrait d'identifier beaucoup plus rapidement l'arrêt de navires en mer, et d'avoir une meilleure compréhension du passé de ces navires.
- Si la phase 3 seule est identifiée (symboles de navires qui se suivent), le contrôleur peut toutefois retracer la route passée récente de ces navires, et retrouver un cas d'approche de deux navires d'origines différentes. L'analyse géovisuelle lui donnerait alors un moyen de visualiser les comportements passés, et par son expérience, il pourrait éventuellement reconnaitre un cas à risque.
- Si deux navires sont identifiés comme suspects dès la phase 2, l'utilisation d'outils classiques tels que le CPA permet déjà d'estimer ou non une approche, et donc un possible scénario de collision ou d'attaque. Une analyse approfondie doit alors être menée sur l'identité des navires et leur dynamique actuelle : l'utilisation de visualisation simple pour reconnaitre le profil dynamique de ces navires permettrait alors d'avoir une meilleure idée du scénario en cours.

De plus, ce même cas d'étude nous permet d'identifier des limites de l'EAAG. A partir de ce même cas d'utilisation (scénario d'attaque et étude en temps réel), il nous est possible de faire varier le profil d'utilisateur en choisissant des compétences relatives aux statistiques et à la visualisation plus fortes. Toutefois, le résultat final des méthodes proposées ne change pas : à cela, nous pouvons trouver plusieurs raisons et solutions pour contrer ces limites.

- Les visualisations qui ont été modélisées dans notre EAAG ne sont pas adaptées au temps réel, mais proposent plus de possibilités pour l'étude de données a posteriori. Il peut donc s'agir d'un cas qu'il n'est **pas possible de reconnaitre en temps réel**, mais seulement par une analyse a posteriori.

- Ce cas peut être reconnu, et dans ce cas notre modèle pourrait être **enrichi en méthodes de visualisation existantes mais non prises en compte ici**, qui soient alors mieux adaptées à l'étude de données en temps réel, notamment lorsque l'on prend en compte un profil d'utilisateur plus avancé sur les aspects techniques
- Les visualisations qui permettent de traiter de telles données, dans le temps imparti, et pour un tel profil d'utilisateur n'existent peut-être pas : cela permet donc de savoir dans quel axe doivent **s'orienter les prochaines propositions et développements de visualisation d'information géographique**.

En effet, une démarche « orientée analyse géovisuelle » aurait alors plus d'intérêt dans le cas de l'analyse des données a posteriori, c'est-à-dire pour le travail des analystes et des chercheurs, afin de comprendre le contexte de l'approche (visibilité, changements de vitesse, etc.) et modéliser les comportements. Dans le modèle que nous proposons, une grande partie des visualisations modélisées concernent en effet le traitement de grandes quantités de données et la reconnaissance d'anomalies au sein de ces quantités de données : cartes de densité composée, coordonnées parallèles, lentille temporelle, carte de Kohonen, etc. Ces méthodes d'analyse de l'information sont en effet plus adaptées à l'étude de grandes quantités de données, a posteriori, plutôt que l'analyse de deux objets mobiles en temps réel.

Par l'utilisation de ces cas d'étude, nous pouvons voir que notre environnement d'aide à l'analyse géovisuelle, en faisant varier les scénarios, la temporalité des données, et le profil des utilisateurs, permet aussi d'identifier ses propres limites : il met en avant les cas où de nouveaux modèles (de visualisations) doivent être proposés, afin de venir compléter et améliorer les réponses du système. Toutefois, l'application de cet EAAG met en avant plus de méthodes pour certains cas d'étude, tels que la caractérisation de l'espace et l'étude de grandes quantités de trajectoires : celles-ci sont proposées lorsque les compétences scientifiques de l'utilisateur varient. Cela permet alors de mettre en avant les **besoins de formation** pour certains utilisateurs de l'information géographique, tels que les contrôleurs, sans nécessairement changer leurs missions d'étude du trafic maritime. En conservant un même cas, mais en augmentant de telles compétences, les autres visualisations proposées pourraient être utilisées : celles-ci nécessitant des statistiques plus avancées ou des modes de représentation (projection et abstraction) de l'information plus complexes.

Dans la prochaine sous-partie de ce chapitre, nous présentons les performances du raisonneur de cet environnement d'aide à l'analyse géovisuelle, mais qui doit être distingué du temps de traitement de l'information pour sa représentation. Comme notre environnement n'est pas lié à un logiciel de représentation spécifique, les méthodes et librairies utilisées pour développer de manière concrètes ces visualisations peuvent changer selon les utilisateurs et leur cadre de travail. Notre application se base par exemple sur l'utilisation de la plateforme web FishEye, pour laquelle nous avons développé certaines des visualisations proposées par notre environnement d'aide à l'analyse géovisuelle. Les performances qui représentent l'utilisation de cet EAAG sont donc spécifiquement celles du raisonneur et de l'ontologie.

4.3. Performance du raisonneur ontologique

Au sein de l'ontologie proposée, nous avons conceptualisé 31 méthodes de visualisation d'informations spatiales, temporelles et sémantiques. Ces méthodes peuvent ainsi être utilisées de manière complémentaire, grâce à un éventuel lien dynamique entre celles-ci. Lors du processus de

raisonnement, le moteur d'inférence est donc amené à lire et classer ces différentes méthodes de visualisation, selon les règles développées et les informations en entrée correspondant au cas d'utilisation. Le graphique présenté sur la Figure 4-25 présente le temps (en ms) utilisé par le raisonneur Pellet, afin de classifier les objets de l'ontologie à partir des règles, selon le nombre de méthodes de visualisation conceptualisées. On remarque que le temps d'exécution adopte une croissance exponentielle en fonction du nombre de visualisations modélisées, qui peut s'expliquer par la capacité itérative du raisonneur choisi. En effet, plus les concepts de l'ontologie sont nombreux, plus les règles vont ajouter et modifier les attributs de ceux-ci, et ainsi nécessiter des boucles de raisonnement supplémentaires.

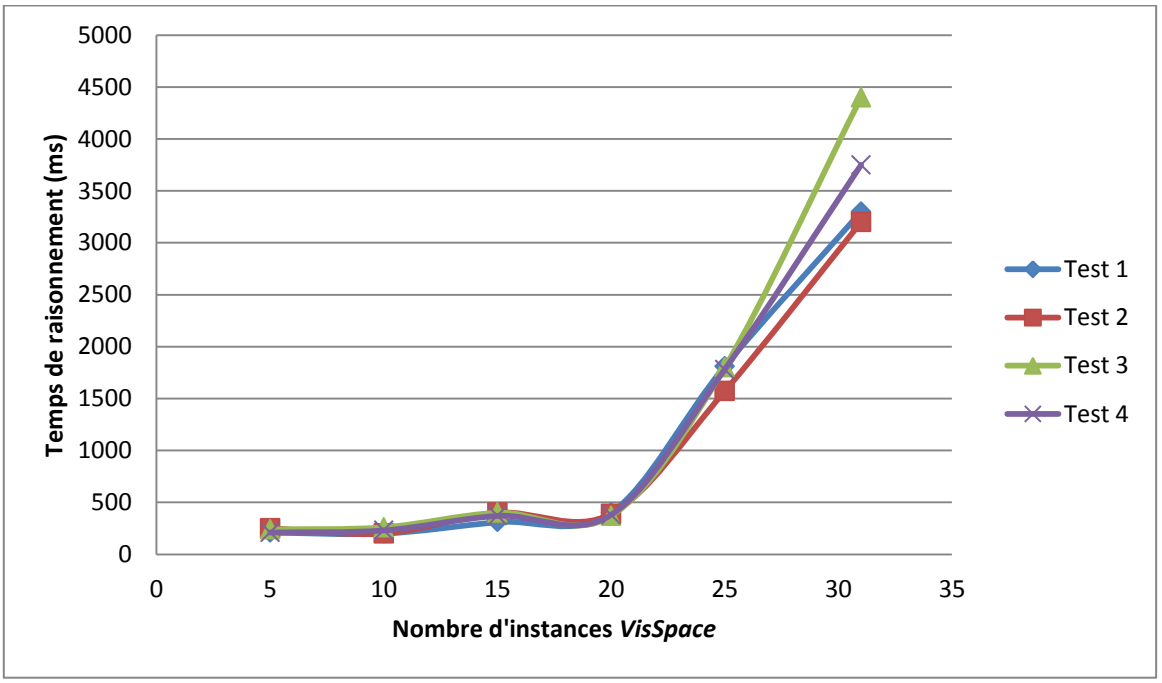


Figure 4-25. Temps (en ms) utilisé par la raisonneur Pellet pour classifier les objets, selon le nombre de méthodes de visualisation modélisées

Pour les 31 méthodes modélisées, le temps d'exécution de l'inférence au sein de l'ontologie reste inférieur à 5 secondes ; ce résultat reste acceptable pour une utilisation rapide au cours de la recherche de visualisations dans un contexte donné. Néanmoins, ce temps de raisonnement ne prend pas en compte le temps d'exécution de ces méthodes et de la présentation de la visualisation, qui peut être bien plus longue. De manière à limiter le temps d'inférence de ce raisonneur, il est de plus important de limiter le nombre de visualisation modélisées, et de ne pas vouloir proposer un catalogue « exhaustif » des méthodes de visualisation et d'analyse : le temps d'exécution pour l'aide à l'analyse géovisuelle serait alors trop long, et l'intérêt d'un tel environnement serait perdu.

Suite aux exemples que nous avons présentés et à la discussion menée sur le rôle de l'EAAG, nous pouvons tirer des conclusions plus larges, à l'échelle de la gestion de risques, en revenant sur les processus présentés dans le chapitre 1 et le rôle d'un tel environnement dans la gestion des risques. Les enseignements tirés de ces exemples d'applications permettent en effet de mieux comprendre le rôle de l'analyse géovisuelle au sein de ce processus complexe.

4.4. Retour sur la gestion des risques

Lors de la présentation de la gestion des risques par des systèmes cognitifs joints, nous avons identifié les différentes phases où l'analyse géovisuelle pouvait jouer un rôle. Nous pouvons reprendre ce tableau afin de dresser le bilan de cette partie de démonstration, en comparant le rôle de l'analyse géovisuelle à une approche plus générale des outils au service de la gestion des risques. Dans les exemples que nous avons présentés, certaines étapes et certains rôles ont souligné le rôle fondamental de l'analyse géovisuelle, au cours de l'exploration et l'analyse des comportements à risques. La Table 4-7 reprend le bilan effectué dans la partie 1.1.1.5, en prenant en compte les résultats de l'analyse géovisuelle appliquée aux risques.

Sur ce tableau, nous reprenons les termes identifiés dans les différents modèles de la communauté des outils pour la gestion des risques, et nous évaluons le rôle effectif que peut jouer l'analyse géovisuelle, du moins efficace (0) au plus efficace (+++). Cette observation est faite selon le nombre méthodes renvoyées par l'EAAG, selon les différentes d'observation et d'analyse. Grâce aux exemples que nous avons présentés, on retrouve un rôle bien plus important de l'analyse géovisuelle dans les phases de modélisation et d'orientation du processus de contrôle des risques, notamment par le pouvoir de **caractérisation des espaces, des objets et de leur dynamique**, qui permettent alors la mise à jour des concepts de risques et de comportements à risque par l'analyse. La littérature du domaine montre néanmoins le rôle important que joue l'analyse géovisuelle dans les domaines de la prise de décision : mais dans ce cas, le public visé est beaucoup plus large et la chaine géodécisionnelle bien plus complexe qu'une simple analyse de données. Dans un cadre de prise de décision, le développement de solutions géodécisionnelles doit prendre en compte le processus de décision, les multiples acteurs à intégrer, les actions sur le terrain et leur compensation en temps réel. Les solutions SOLAP sont un bon exemple d'environnement d'analyse géovisuelle adapté au contexte décisionnel.

Enfin, par les limites du traitement et de la représentation de données en temps réel, notamment dus aux biais cognitifs humain, l'analyse géovisuelle reste un domaine « trop haut » pour les phases de simple observation d'un système contrôlé. Toutefois, elle peut présenter des interfaces de grand intérêt, modulo la formation des utilisateurs.

Table 4-7. Bilan du rôle de l'analyse géovisuelle dans un processus de gestion des risques

Etape / Modèle	Anticipation	.....		Compensation
OODA (années 60, Boyd 1996)	Observer +	Orienter +++	Décider +	Agir 0
COCOM (Hollnagel 1993)	Evènements système +	Mise à jour concepts +++		Action 0
Contrôle système socio-technique (Hollnagel 2004)		Caractérisation du contexte +++ Identification des risques		

ECOM (Hollnagel et Woods 2005)		++ Analyser les risques +++ Identification des contremesures +		
	Surveillance +	Définition but ++	Régulation 0	Suivi +

Grâce à ce retour sur la gestion des risques, en général, nous pouvons ainsi mieux situer la force de notre travail au sein de ce processus. Notre proposition d'un environnement d'aide à l'analyse géovisuelle est surtout très adaptée à des acteurs tels que les analystes et les personnes amenés à étudier des données de trafic passé. Dans le cas de l'observation d'informations en temps réel dans le processus de contrôle et de gestion des risques, notre environnement nécessite de prendre en compte des méthodes supplémentaires afin d'affirmer son rôle du point de vue de l'urgence immédiate ; il vient aussi appuyer l'utilité de former des utilisateurs à une visualisation plus avancée que des outils de base, de manière à pouvoir prendre en compte la visualisation des trajectoires à la fois dans l'espace et le temps. Ces méthodes existent, peuvent être utilisées pour la visualisation d'informations en temps réel, mais nécessite un changement des habitudes des utilisateurs concernés, comme ont pu le montrer des recherches récentes (Davenport et Risley 2006 ; van Laere et Nilsson 2009 ; Riveiro 2011).

Le rôle d'un tel environnement d'aide à l'analyse géovisuelle, intégré au processus de gestion des risques, montre donc une contribution majeure : l'utilisation optimale de méthodes visuelles pour la compréhension du trafic maritime, et ainsi l'amélioration de l'anticipation et de la reconnaissance future des menaces en mer.

4.5. Perspective d'amélioration du module

4.5.1. Ordre de proposition des visualisations

L'interface développée permet actuellement de stocker les informations sur l'utilisation de l'environnement d'aide à l'analyse géovisuelle. Nous enregistrons ainsi les informations données en entrée (profil de l'utilisateur, questions, etc.) ainsi que les visualisations qui sont choisies à la fin du processus. Ces informations sont envoyées sur un serveur, lors de la validation des choix de visualisation par l'interface Java. De cette manière, un fichier XML permet de tracer l'utilisation effective de cet environnement.

Par la suite, une amélioration possible de cet environnement serait de prendre en compte les résultats enregistrés dans ce fichier, afin de faire évoluer dynamiquement l'ontologie. Pour cela, deux propriétés supplémentaires peuvent être utilisées dans l'ontologie pour qualifier l'utilisation des différentes méthodes de visualisation : le nombre de sélections, ou d'utilisations si

nous supposons que chaque méthode sélectionnée est effectivement utilisée par l'opérateur ; ainsi qu'un indice de confiance pour cette utilisation, qui représente un pourcentage proportionnel au nombre d'utilisations.

Par défaut, l'ontologie accorde un « poids » équivalent à toutes les méthodes de visualisation modélisées. Grâce à ces variables, il serait ainsi possible de modifier ce poids au cours des utilisations, en vue d'ordonner les méthodes proposées dans l'interface : la première visualisation serait ainsi la plus choisie, et la dernière celle qui est utilisée moins fréquemment.

4.5.2. Auto-évaluation du profil utilisateur

Tel que nous l'avons proposé dans cette version de l'EAAG, le profil utilisateur dépend de trois variables qualitatives ordonnées (les compétences) et une variable qualitative nominale (environnement de travail). En proposant cette modélisation de l'utilisateur, orienté pour développer des règles d'utilisation, nous avons proposé des rôles-types de façon à donner des exemples de valeurs à ces paramètres. Toutefois, lors de l'utilisation de l'interface de sélection de visualisation, un utilisateur quelconque ne peut pas juger de lui-même des paramètres à rentrer. Les valeurs de ces échelles Faible-Moyen-Avancé représente des évaluations empiriques des compétences de chacun.

Une amélioration importante pour l'utilisabilité de cet outil serait donc de proposer une grille d'évaluation pour les utilisateurs, de manière à savoir quelles valeurs utiliser pour compléter ce profil. En renseignant un maximum d'informations sur leur profil, leurs études, leurs missions et leurs capacités, une étude comparative des différents utilisateurs pourraient permettre d'évaluer l'odre d'importance et l'influence des caractéristiques indeifiées sur le travail d'analyse visuelle. Ainsi, les rôles-types et les valeurs de ces propriétés pourraient dépendre des réponses de l'utilisateur, au cours d'une auto-évaluation de son expérience du domaine, de ses compétences technologiques et de sa connaissance scientifique et statistique.

Conclusion du Chapitre 4

Dans ce chapitre final, nous avons présenté le développement d'une preuve de concept et son intégration dans une plateforme de visualisation et d'analyse par des techniques web. L'utilisation d'une ontologie, orientée web sémantique, justifie le développement de cette application web et prouve la faisabilité d'un tel environnement d'aide à l'analyse. Dans cet outil web, nous avons pris en compte certaines des visualisations modélisées dans la base de connaissances, afin d'en démontrer l'utilisation par la suite.

Nous sommes alors partis de plusieurs scénarios, relatifs à différents risques maritimes, qui peuvent être qualifiés de « complexes » si l'on se base sur les fonctionnalités actuelles des systèmes de surveillance maritime. Par complexe, nous signifions qu'un comportement, ses causes et ses conséquences, ne peuvent être identifiées que par une analyse avancée des informations. Actuellement, les systèmes de surveillance ne permettent d'identifier que des évènements simples ou compliqués, si l'on reprend la classification Cynefin. Ces derniers demandent une certaine habitude de manipulation de ces outils, ainsi qu'une connaissance des données et comportements « habituels ». Toutefois, lorsqu'un comportement est rencontré de manière plus rare, il peut être caractérisé comme complexe. Il n'est en effet pas reconnu aussi rapidement que les évènements habituels.

A partir de quatre scénarios, qui peuvent avoir une représentation cartographique presque similaire, nous avons différencié les évènements qui les identifient, ainsi que les tâches d'analyse qui permettent de reconnaitre ces évènements. Nous avons alors développé ces différents scénarios, ainsi que les utilisations comparées des systèmes de surveillance actuels et d'un système assisté par l'EAAG que nous proposons. De ce cas d'utilisation, nous avons identifié à la fois les atouts et les limites de notre proposition d'aide à l'analyse géovisuelle. Un tel environnement, comme nous l'avons proposé en preuve de concept, permet à la fois de proposer des méthodes non exploitées actuellement grâce à une sélection de méthodes de visualisation, mais aussi d'identifier les étapes où de nouvelles visualisations **doivent être modélisées** (dans notre ontologie) ou bien **proposées par la communauté scientifique**, si celles-ci n'existent pas. Outre le support à l'analyse géovisuelle, qui serait entièrement fonctionnel si beaucoup plus de méthodes étaient prises en compte, il est donc intéressant de voir cette proposition comme un catalogue des possibilités et limites de l'analyse géovisuelle, qui guiderait en plus le développement de nouvelles visualisations.

Pour cela, comme nous l'avons vu au cours des perspectives d'amélioration de cette interface, l'étude de l'utilisation effective d'un tel système, par les rapports d'utilisation et d'analyse, pourrait être prise en compte pour les propositions futures de la communauté GeoVis, en termes d'analyse de mouvements.

CONCLUSION ET PERSPECTIVES

Les travaux présentés dans ce manuscrit de thèse prennent à la fois part aux recherches dans les domaines des sciences de l'information géographique, de la gestion de connaissances, tout en prenant une problématique opérationnelle sur l'utilisation de la visualisation. Les problèmes opérationnels identifiés provenaient en parties de questions scientifiques qui ont pu guider notre recherche et notre travail de modélisation. Le terrain d'étude qu'est la sécurité et la sûreté de la maritimisation de l'énergie, a permis de poser un cadre applicatif pour des questions théoriques, comme support de travail et de démonstration. L'objectif de cette recherche était de proposer une modélisation des connaissances en visualisation de l'information spatio-temporelle telle que des données de mouvement, pour l'analyse des comportements des navires.

Les étapes clés de cette recherche ont été d'étudier l'usage de l'analyse géovisuelle pour l'étude des mouvements d'objets mobiles, de formaliser ces solutions et ainsi définir et proposer un environnement d'aide à l'analyse géovisuelle. Pour cela, nous avons choisi de fonder notre travail sur l'utilisation d'ontologies. Ce choix de modélisation permet de décrire différents concepts qui peuvent être reliés, complété par des règles de raisonnements, et partagés par un formalisme reconnu par le web sémantique. Pour cela, nous avons proposé une ontologie relative au domaine d'application, afin de décrire les données et les risques maritimes, ainsi que le mode d'analyse des comportements de navires. De manière parallèle, nous avons proposé une modélisation du processus de visualisation d'information spatiale, temporelle et sémantique, de manière à supporter l'analyse de ces comportements.

De telles méthodes de visualisation de l'information géographique ont pu être intégrées à une plateforme web pour la visualisation et l'analyse du trafic maritime, de façon à tester certaines méthodes formalisée dans un environnement complet. L'intégration du modèle ontologique au sein d'un module d'aide à l'analyse géovisuelle a été illustré, grâce à une interface de sélection de méthodes de visualisation. Une telle interface illustre le rôle de l'aide à l'analyse géovisuelle, qui pourrait être intégrée dans les nombreux environnements d'analyse géovisuelle proposés par la communauté scientifique. Cela aurait pour but de faire découvrir des visualisations que l'utilisateur n'a pas pour habitude d'utiliser, tout en respectant ses tâches à accomplir et ses limites.

Bilan de la thèse

Au commencement de cette recherche, l'idée était de proposer une nouvelle manière de visualiser le trafic maritime, afin de mieux comprendre les comportements et ainsi mieux gérer les risques relatifs au trafic maritime. Toutefois, au vu des nombreux travaux de la communauté des sciences de l'information géographique, nous avons identifié la limite qu'est **le manque de formalisation des méthodes d'analyse géovisuelle**, ainsi que la **non-utilisation opérationnelle des visualisations proposées par la recherche**. L'être humain étant au cœur de toute prise de décision, nous avons choisi d'orienter notre recherche sur une **aide à l'analyse géovisuelle**, et non pas une automatisation du processus d'analyse des mouvements.

Définition de l'aide à l'analyse géovisuelle pour les utilisateurs

Guidés par les courants des systèmes d'aide à la décision et des environnements de résolution de problèmes, nous avons défini ce que nous appelons un **environnement d'aide à l'analyse géovisuelle**. Un tel type d'environnement ne prend pas des décisions à la place de l'homme, mais permet de soutenir l'analyse d'informations par l'utilisation de méthodes visuelles adaptées au cas d'utilisation. Nous avons identifié plusieurs profils d'utilisateurs qui auraient besoin de ce type d'environnement, à savoir les personnes liés à la prise de décision à partir de l'analyse de données de mouvement (les contrôleurs, les analystes, etc.), mais aussi les scientifiques amenés à analyser l'information géographique afin de modéliser les risques maritimes. Ces nombreux utilisateurs, par leur profil, leur formation et leurs habitudes, sont amenés à utiliser des visualisations variées, dans des contextes différents. Aussi, un outil d'aide au choix de visualisation pour la stratégie d'analyse visuelle peut venir supporter leur travail quotidien.

Conception d'un modèle pour la formalisation de la visualisation

Au-delà de l'apport « opérationnel » de cet environnement, nous avons proposé un modèle pour la formalisation des méthodes de visualisation, qui sont interprétées comme la projection d'information (concrète ou non) vers un espace visuel. Comme nous l'avons le présenté dans ce manuscrit, de nombreux travaux ont proposé des modèles reliés à un langage ou bien un outil particulier, afin d'automatiser la visualisation d'information. Toutefois, ces travaux sont difficilement **accessibles et partageables**. L'utilisation d'un formalisme tel que les ontologies nous a permis de proposer un modèle, publié en ligne, qui peut alors être partagé, repris, amélioré, et donc maintenu par les communautés intéressées. Grâce à ce type de modèle, la diffusion de méthodes de visualisation et leur utilisation peut donc être améliorée, et faire avancer à la fois les travaux des personnes liées au développement de telles méthodes, et à leur utilisation opérationnelle.

Intégration dans un environnement d'analyse du trafic maritime

Au sein de l'outil web FishEye, nous avons intégré plusieurs méthodes de visualisation d'informations spatiales, temporelles et sémantiques, pour l'exploration et l'analyse du trafic maritime. Cet environnement se base sur plusieurs travaux de recherche, menés précédemment dans notre centre de recherche. Cet outil permet donc d'intégrer de nombreux modèles, algorithmes et représentation de la dynamique des navires, et l'étude de ces comportements. Pour illustrer l'utilisation de la formalisation des solutions d'analyse géovisuelle, notre prototype a pu être appliqué dans cette architecture web, afin de permettre le raisonnement sur des cas concrets de risques maritimes.

Ouverture à la résilience

Tel que nous l'avons présenté dans un article de conférence (Vatin et al. 2014), notre proposition peut avoir un rôle d'intérêt dans la définition d'un système résilient, afin d'appuyer sa place dans la gestion des risques émergents au sein de systèmes « complexes » (Vatin et Napoli 2013f). La notion de résilience n'a pas été au cœur de notre sujet de recherche, mais elle s'est avérée être une ouverture intéressante, qui pourrait alors prendre part dans les axes de recherche de la communauté de l'ingénierie de la résilience (Comfort et al. 2010 ; Hollnagel et al. 2010). Dans ce cadre, nous pouvons considérer la résilience d'un système socio-technique comme étant la capacité du système à **adopter le bon comportement face à une situation connue anticipée**, **connue mais non anticipée** ou bien **face à une surprise** : afin de répondre de manière adaptée, le système doit **détecter** et

identifier un évènement, **évaluer** sa gravité et **répondre de manière efficace** dans le temps imparti (Hollnagel et al. 2010). L'EAAG que nous proposons permet d'identifier des méthodes de visualisations qui soient adaptées à un contexte, nouveau ou déjà connu, afin de faciliter la compréhension des informations. Ainsi, un tel système participe à l'amélioration de la prise de décision par une assistance à l'analyse, et **accordant ainsi plus de temps à l'analyse qui doit être menée**. Selon la définition donnée, nous pouvons considérer qu'un tel EAAG participe, de manière indirecte, à rendre un système plus résilient : ceci est une piste de recherches et d'approfondissement tout à fait intéressante pour notre centre de recherche, d'un point de vue social et technique.

Perspectives

Suite à notre travail de modélisation, de développement et d'intégration d'un tel environnement d'aide à l'analyse géovisuelle, nous pouvons identifier plusieurs perspectives majeures pour l'amélioration de ces différentes étapes. Nous présentons ainsi les perspectives sur le modèle choisi, puis sur l'utilisation opérationnelle. Enfin, nous ouvrons sur des perspectives plus larges concernant l'analyse géovisuelle au sein d'un système tel que l'ensemble des acteurs de la sécurité et sûreté de la maritimisation de l'énergie.

Perspectives de modélisation

Un profil utilisateur et un environnement de travail plus complexe

Comme nous l'avons expliqué à la fin du chapitre 4, depuis l'interface que nous avons proposé, il est encore compliqué et trop subjectif pour un utilisateur de pouvoir évaluer son profil sur une telle échelle empirique, selon les trois niveaux de compétence et d'expérience identifiés dans le chapitre 3. Une perspective importante pour ce travail de recherche serait d'approfondir la modélisation de l'utilisateur et de son environnement, afin de rendre le profil plus intuitif, et moins empirique. Notre preuve de concept illustre les effets de sélection selon les changements de ces valeurs caractéristiques, mais ne reflètent pas une réalité concrète et quantifiable. Ce modèle utilisateur peut donc subir de grandes améliorations afin de représenter au mieux la réalité, et un panel plus important d'utilisateurs potentiels. Le contexte d'utilisation pourrait aussi être approfondi : dans notre travail de modélisation, nous n'avons utilisé qu'une valeur nominale ayant pour seul attribut sa mobilité ou non. Plusieurs travaux de modélisation de l'utilisateur et de son environnement ont pu être proposés, cette question étant un point de recherche à part entière ; tandis que nous avons préféré exploiter la modélisation de la visualisation elle-même.

Des étapes d'analyse chronologiques dans la stratégie de visualisation

La modélisation des méthodes de visualisation d'information que nous avons proposée ne prend pas en compte la dynamique du processus d'analyse et de décision. Pourtant, comme le montre la boucle OODA et les différents processus de gestion des risques, l'analyse et la définition des comportements à risque doit être un processus itératif, tout comme la recherche d'information dans un tel système. Au cours de l'exemple du scénario d'attaque, déroulé dans le chapitre 4, nous avons constaté que certaines tâches sont constamment nécessaires pendant l'intégralité du processus d'exploration et d'analyse des données, tandis que d'autres tâches n'interviennent qu'à une unique étape dans l'analyse du comportement et l'identification du risque. Une amélioration d'intérêt serait de prendre en compte l'ordre et la récurrence de ces tâches d'analyse géovisuelle dans la modélisation des

stratégies de visualisation, et d'avoir un impact sur les résultats proposés : ordre des propositions, méthode la plus importante car plus la demandée, etc. Durant nos travaux, nous avons tenté de modéliser les concepts *Goal* par une suite ordonnée de *Task*, mais le modèle final était beaucoup plus compliqué à gérer et mettre à jour, notamment lors du développement des règles. Nous sommes donc revenus à une version non chronologique des tâches d'exploration, mais en essayant d'être le plus exhaustif possible.

Perspectives d'application

Ouverture à l'analyse générale du mouvement

Cette étude a été menée à partir d'un corpus de textes et d'études de nombreux domaines différents, tout en concernant essentiellement l'analyse de données de mouvement. Ces études traitent entre autres de mouvements de navires (sécurité maritime), d'avions (sécurité aérienne), de voitures ou encore de piétons et cyclistes (études de transport). Le domaine du maritime ayant essentiellement été utilisé dans la mise en place de la modélisation des données et des tâches métiers, un tel environnement d'aide à l'analyse géovisuelle pourrait être appliqué à ces autres domaines en modifiant les objets étudiés et les spécificités des ces données (tâches métiers, espace géographique, etc.). L'étude de données routières prend en compte des données importantes mais presque inexistantes dans le domaine maritime : la contrainte du réseau routier pour le déplacement des objets. Cette contrainte permet notamment de simplifier certaines représentations cartographiques, en passant d'un espace ouvert à un espace de liens et de nœuds.

Toutefois, l'approche proposée dans cette thèse permet d'ouvrir le domaine de l'étude des mouvements d'objets mobiles, quel que soit l'espace concerné, à de nombreuses méthodes géovisuelle pour la représentation, l'exploration et l'analyse de données.

Comparaison avec une approche SOLAP

Nous avons brièvement présenté le concept de *Spatial Online Analytical Process* (SOLAP) lors du chapitre 2. Il est intéressant de revenir sur cette solution d'analyse de l'information géographique, maintenant que notre approche par un environnement d'aide à l'analyse géovisuelle a pu être présentée. Le SOLAP a en effet été proposé par la communauté afin de proposer une nouvelle façon de visualiser et d'analyser la donnée spatiale, dans une problématique de prise de décision. Cette nouvelle solution se base sur une architecture de données très spécifique (en étoile, en flocon), permettant ainsi un calcul préliminaire des données d'analyse à différentes échelles spatiales, temporelles et sémantiques. Une interface SOLAP permet donc de manipuler l'information dans un maximum de dimension, grâce aux nombreuses fonctionnalités de pivot, forage vers le bas ou vers le haut, et forage latéral.

Une interface SOLAP, telle que l'outil Map4Decision, se présente comme l'utilisation de nombreuses visualisations par des cartes, des graphes et des diagrammes, afin de prendre en compte la pluralité des dimensions de l'information analysée. Pour cela, nous pouvons considérer une interface SOLAP comme un environnement d'analyse géovisuelle à part entière, selon la définition que nous avons choisie pour ce terme. Toutefois, sa spécificité réside en l'architecture des données, leur mode de traitement et les manipulations possibles dans une solution (S-)OLAP.

Notre proposition d'un environnement d'aide à l'analyse géovisuelle ne prend pas en compte l'architecture des données étudiées, celles-ci étant considérée de manière indépendantes les unes des autres, en vue de leur projection dans les espaces de visualisation. Il serait alors relevant de tester l'usabilité d'une telle interface de sélection de visualisation dans une optique SOLAP, c'est-à-dire afin d'assister l'utilisateur dans son usage d'environnement SOLAP lors de l'étude de données. Les travaux de Beaulieu (Beaulieu 2009) se sont notamment inscrit dans ce cadre d'aide à l'utilisation de SOLAP, en proposant une stratégie de visualisation adaptée aux données d'entrée et leur structure. Ce travail portait essentiellement sur l'utilisation des règles de Bertin afin de mieux structurer et présenter l'information au sein d'interface SOLAP.

Le couplage d'un tel environnement interactif pour l'exploration de données spatiales, temporelle et sémantique, à un environnement d'aide à l'analyse géovisuelle, serait alors un apport intéressant pour la communauté des sciences de l'information géographique, dans l'optique d'améliorer l'usabilité des solutions d'analyse géovisuelle. Toutefois, tel que nous avons présenté ce prototype, notre environnement est très lié à l'analyse de flux et de mouvements tels que les déplacements de navires, tandis que l'utilisation de SOLAP est mieux adaptée à l'étude de données géographiques tels que des unités spatiales, démographiques, etc.

Recherches sur la qualité de la décision et de l'analyse visuelle

La philosophie de **conception centrée utilisateur**, définie par la norme ISO 9241-210 (International Standards Organisation 2010), est un mode de conception d'interfaces homme-machine, qui met en avant l'implication de l'utilisateur final dans cette étape de développement. L'une des principales finalités de cette démarche est d'améliorer l'ergonomie de l'interface proposée. Dans le cadre de notre proposition, l'interface a été développée comme preuve de concepts pour démontrer l'utilisation de notre base de connaissances, sans intégrer l'utilisateur dans le développement de cet outil. Toutefois, l'ergonomie de l'interface finale est un facteur clé dans son utilisation effective (voir la partie de cette thèse sur la facilité d'utilisation perçue, de Davis). Présenter des visualisations adaptées, mais d'une manière peu ergonomique, ne motive pas nécessairement à leur utilisation. De plus, si l'on prend en compte de nombreuses visualisations sélectionnées par l'environnement d'aide à l'analyse géovisuelle, « plus de visualisations » ne signifie pas nécessairement « une décision plus rapide et efficace », mais peut-être une **compréhension moins complète de la situation**.

En plus de la question de l'ergonomie, se pose alors la question de la qualité de la décision qui découlera de l'utilisation de ces environnements d'analyse. Comment qualifier, en plus de la qualité des données utilisées, la qualité des conclusions tirées de la visualisation d'information ? Cette piste de recherche doit prendre en compte de nombreux facteurs, tels que :

- la qualité des données (ISO 19157:2013), dont nous n'avons pas parlé dans cette thèse, mais qui est fondamentale pour la prise de décision (précision de l'enregistrement, résolution de l'information, etc.)
- le « niveau d'abstraction » des données représentées, qui donne alors lieu à une qualité de la représentation visuelle (précision spatiale, sémantique, temporelle)
- le nombre de données manquantes afin d'avoir un scénario validé à 100% (nombre d'évènements remarquables, valeurs des attributs, etc.) et donc une décision sûre

La question se pose donc, d'un point de vue opérationnel : à quel niveau de « sûreté », dans le sens confiance, un opérateur va-t-il prendre une décision et ordonner une opération sur le terrain, si besoin est ? Ces nombreuses questions de recherches doivent alors faire intervenir des notions de qualité de l'information, de qualité des systèmes d'information, et de sociologie.

BIBLIOGRAPHIE

Abdullah, Ahsan. 2009. « Analysis of mealybug incidence on the cotton crop using ADSS-OLAP (Online Analytical Processing) tool ». *Computers and Electronics in Agriculture* 69 (1): 59‑72.

Adler, Richard M., et Jeff Fuller. 2007. « An integrated framework for assessing and mitigating risks to maritime critical infrastructure ». In *2007 IEEE Conference on Technologies for Homeland Security*, 252‑57.

Ahola, Terhi, Kirsi Virrantaus, Jukka Matthias Krisp, et Gary J. Hunter. 2007. « A spatio-temporal population model to support risk assessment and damage analysis for decision-making ». *International Journal of Geographical Information Science* 21 (8): 935‑53.

Akella, Mamata Kumari. 2009. « First responders and crisis map symbols: Clarifying communication ». *Cartography and Geographic Information Science*, Human-computer interaction handbook, 36 (1): 19‑28.

Andrienko, Gennady, et Natalia Andrienko. 1999. « Making a GIS intelligent: CommonGIS project view ». In *Proceedings of AGILE Conference*, 19‑24. Rome, Italy: 04/1999.

Andrienko, Gennady, et Natalia Andrienko. 2001. « Constructing parallel coordinates plot for problem solving ». In *1st International Symposium on Smart Graphics*, 9‑14. http://geoanalytics.net/and/papers/sg01.pdf.

Andrienko, Gennady, et Natalia Andrienko. 2009. « Interactive spatio-temporal cluster analysis of VAST Challenge 2008 datasets ». In *Proceedings of the ACM SIGKDD Workshop on Visual Analytics and Knowledge Discovery: Integrating Automated Analysis with Interactive Exploration*, 5‑11. Paris, France.

Andrienko, Gennady, et Natalia Andrienko. 2010. « A general framework for using aggregation in visual exploration of movement data ». *The Cartographic Journal* 47 (1): 22‑40.

Andrienko, Gennady, Natalia Andrienko, Urska Demšar, Doris Dransch, Jason Dykes, Sara Irina Fabrikant, Mikael Jern, Menno-Jan Kraak, Heidrun Schumann, et Christian Tominski. 2010. « Space, time and visual analytics ». *International Journal of Geographical Information Science* 24 (10): 1577‑1600.

Andrienko, Gennady, Natalia Andrienko, Christohe Hurter, Salvatore Rinzivillo, et Stefan Wrobel. 2011. « From movement tracks through events to places: Extracting and characterizing significant places from mobility data ». In *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 161‑70.

Andrienko, Gennady, Natalia Andrienko, Piotr Jankowski, Daniel Keim, Menno-Jan Kraak, Alan M. MacEachren, et Stefan Wrobel. 2007. « Geovisual analytics for spatial decision support: Setting the research agenda ». *International Journal of Geographical Information Science* 21 (8): 839‑57.

Andrienko, Natalia, et Gennady Andrienko. 2007. « Intelligent Visualisation and Information Presentation for Civil Crisis Management ». *Transactions in GIS* 11 (6): 889‑909.

Andrienko, Natalia, et Gennady Andrienko. 2011. « Spatial generalization and aggregation of massive movement data ». *IEEE Transactions on Visualization and Computer Graphics* 17 (2): 205‑19.

Andrienko, Natalia, et Gennady Andrienko. 2013. « Visual analytics of movement: an overview of methods, tools and procedures ». *Information Visualization* 12 (1): 3‑24.

Andrienko, Natalia, Gennady Andrienko, et Peter Gatalsky. 2003. « Exploratory spatio-temporal visualization: an analytical review ». *Journal of Visual Languages & Computing* 14 (6): 503‑41.

Andrienko, Natalia, Gennady Andrienko, et Peter Gatalsky. 2005. « Impact of Data and Task Characteristics on Design of Spatio-Temporal Data Visualization Tools ». In *Exploring*

Geovisualization, édité par Jason Dykes, Alan M. MacEachren, et Menno-Jan Kraak, 201‑22. Elsevier Ltd.

Andrienko, Natalia, Gennady Andrienko, Hans Voss, Fatima Bernardo, Joana Hipolito, et Ursula Kretchmer. 2002. « Testing the usability of interactive maps in CommonGIS ». *Cartography and Geographic Information Science* 29 (4): 325‑42.

Antoni, Jean-Philippe, Olivier Klein, et Stéphane Moisy. 2004. « Cartographie interactive et multimédia : vers une aide à la réflexion géographique ». *Cybergeo : European Journal of Geography*, novembre.

Antoniou, Grigoris, et Frank van Harmelen. 2009. « Web Ontology Language: OWL ». In *Handbook on ontologies*, édité par Steffen Staab et Rudi Studer, 2nd éd., 67‑92. Berlin Heidelberg: Springer Verlag.

Bachimont, Bruno. 2000. « Engagement sémantique et engagement ontologique: conception et réalisation d'ontologies en ingénierie des connaissances ». In *Ingénierie des Connaissances : Évolutions récentes et nouveaux défis*, édité par Jean Charlet, Manuel Zackland, Gilles Kassel, et Didier Bourigault, 305‑23. Eyrolles.

Bagozzi, Richard P. 2007. « The legacy of the technology Acceptance Model and a proposal for a paradigm shift. » *Journal of the Association for Information Systems* 8 (4): 244‑54.

Balzarini, Rafaella, Paule-Annick Davoine, et Muriel Ney. 2013. « Processus cognitifs experts vs étudiants lors de la réalisation de cartes de susceptibilité ». In *Actes de la Conférence Internationale de Géomatique et d'Analyse Spatiale 2013*, 87‑99. Brest, France.

Bass, Ellen J., Leigh A. Baumgart, et Kathryn Klein Shepley. 2013. « The effect of information analysis automation display content on human judgment performance in noisy environments ». *Journal of Cognitive Engineering and Decision Making* 7 (1): 49‑65.

Bass, Ellen J., et Amy R. Pritchett. 2008. « Human-automated judge learning: a methodology for examining human interaction with information analysis automation ». *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* 38 (4): 759‑76.

Baudel, Thomas. 2002. « Visualisations compactes : une approche déclarative pour la visualisation d'information ». In *Proceedings of the 14th French-speaking Conference on Human-computer Interaction*, 161‑68. New York, NY, USA: ACM.

Baudel, Thomas. 2004. « Browsing through an information visualization design space ». In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, 765‑66. New York, NY, USA: ACM.

Beaulieu, Véronique. 2009. « Étude de la visualisation géographique dans un environnement d'exploration interactive de données géodécisionnelles - Adaptation et améliorations ». Université de Laval.

Bechhofer, Sean. 2002. *Ontology Language Standardisation Efforts*. Deliverable 4.0. University of Manchester.

Bechhofer, Sean, Carole Groble, et Ian Horrocks. 2002. *Requirements of Ontology Languages*. Deliverable 4.1. University of Manchester.

Bédard, Yvan. 2010. « Le géodécisionnel: origine, évolution, état de l'art, enjeux, R&D ». Sophia Antipolis, France, octobre 29. http://yvanbedard.scg.ulaval.ca/wp-content/documents/slideshow/publication/593/593.pdf.

Bédard, Yvan, Sonia Rivest, et Marie-Josée Proulx. 2007. « Spatial On-Line Analytical Processing (SOLAP): Concepts, architectures and solutions from a geomatics engineering perspective ». In *Data Warehouses and OLAP: Concepts, Architecture and Solutions*, édité par Robert Wrembel et Christian Koncilia, 298‑319. EDA Group.

Béguin, Michèle, et Denise Pumain. 1994. *La représentation des données géographiques : Statistique et cartographie*. Cursus. Paris: Armand Colin.

Belhedi, Amor. 1993. « L'espace géographique : de l'absolu au relatif ». In *L'espace : Concepts et approches*, FSHS, 11‑36. Tunis.

Benbasat, Izak, et Henri Barki. 2007. « Quo vadis, TAM? ». *Journal of the Association for Information Systems* 8 (4): 211‑18.

Bendaoud, Rokia, Mohamed Rouane Hacene, Yannick Toussaint, Bertrand Delecroix, et Amedeo Napoli. 2007. « Construction d'une ontologie à partir d'un corpus de textes avec l'ACF ». In *IC 2007*. Grenoble.

Benkert, Marc, Joachim Gudmundsson, Florian Hübner, et Thomas Wolle. 2008. « Reporting flock patterns ». *Computational Geometry* 41 (3): 111‑25.

Berson, Alex, et Stephen J. Smith. 1997. *Data Warehousing, Data Mining, and Olap*. 1st éd. New York, NY, USA: McGraw-Hill, Inc.

Bertin, Jacques. 1967. *Sémiologie Graphique*. Paris: Mouton/Gauthier-Villars.

Bouejla, Amal, Xavier Chaze, Franck Guarnieri, et Aldo Napoli. 2014. « A Bayesian network to manage risks of maritime piracy against offshore oil fields ». *Safety Science* 68 (octobre): 222‑30.

Bouejla, Amal, Xavier Chaze, Aldo Napoli, Franck Guarnieri, et Benjamin Alhadef. 2012. « Contribution des réseaux bayésiens à la gestion du risque de piraterie contre les champs pétroliers ». In *Proceedings of Workshop Interdisciplinaire sur la Sécurité Globale*. Troyes, France.

Boyd, John R. 1996. « The Essence of Winning and Losing ». http://tobeortodo.com/wp-content/uploads/2011/11/essence_of_winning_losing.pdf.

Boy, Guy. 2011. « A Human-Centered design Approach ». In *The Handbook of Human-Machine Interaction: A Human-Centered Design Approach*, édité par Guy Boy, 1‑20. Farnham, Surrey, England ; Burlington, VT: Ashgate Publishing Limited.

Brisson, Laurent, Martine Collard, et Nicolas Pasquier. 2006. « Ontologie et base de connaissances pour le pré-traitement et post-traitement en fouille de données ». In *Atelier Fouille de Données Complexe de la conférence EGC sur l'Extraction et la Gestion des Connaissances*, 13‑26. Lyon, France.

Brown, Michael, Sarah Sharples, Jenny Harding, Christopher J. Parker, Nick Bearman, Martin Maguire, David Forrest, Muki Haklay, et Mike Jackson. 2013. « Usability of geographic information: current challenges and future directions ». *Applied Ergonomics* 44 (6): 855‑65.

Brunet, Roger. 1987. *La carte, mode d'emploi*. Montpellier; Paris: Fayard.

Buard, Élodie, et Mickaël Brasebin. 2012. « Exploration visuelle de trajectoires de grands animaux ». *Cartes & géomatique*, n° 211: 101‑13.

Cabena, Peter, Pablo Hadjnian, Rolf Stadler, Jaap Verhees, et Alessandro Zanasi. 1997. *Discovering Data Mining: From Concept to Implementation*. Upper Saddle River, NJ: Prentice Hall.

Cai, Guoray. 2005. « Extending distributed GIS to support geo-collaborative crisis management ». *Geographic Information Science* 11 (1): 4‑14.

Card, Stuart K., et Jock Mackinlay. 1997. « The structure of the information visualization design space ». In *Proceedings of the 1997 IEEE Symposium on Information Visualization*, 92‑99. InfoVis '97. Washington, DC, USA: IEEE Computer Society.

Card, Stuart K., Jock Mackinlay, et Ben Shneiderman. 1999. *Readings in Information Visualization: Using Vision to Think*. San Francisco, Califomia: Morgan Kaufmann Publishers.

Cario, Matthieu, et Jean-Eude Gravière. 2013. *Modélisation de comportements de navires à risque à l'aide d'UML*. Brest, France: Ecole Navale.

Caron, Claude, et Stéphane Roche. 2004. *Aspects Organisationnels des SIG*. Paris: Lavoisier.

Chalandon, Xavier. 2003. « Situation Awareness en conception système ». In *Actes des Deuxiemes Journées d'étude en Psychologie ergonomique*, édité par J. M. C. Bastien, 55‑62. Boulogne-Billancourt, Institut de Psychologie, France.

Chatzis, Konstantinos. 1993. « La régulation des systèmes socio-techniques sur la longue durée : le cas du système d'assainissement urbain ». Ecole Nationale des Ponts et Chaussées.

Chaudhuri, Surajit, et Umeshwar Dayal. 1997. « An Overview of Data Warehousing and OLAP Technology ». *SIGMOD Rec.* 26 (1): 65‑74.

Chaze, Xavier, Amal Bouejla, Franck Guarnieri, et Aldo Napoli. 2013. « Causal Probabilistic Modeling with Bayesian Networks to Combat the Risk of Piracy Against Offshore Oil Platforms ». *The Radio Science Bulletin* 345 (Disaster Management special issue).

Chaze, Xavier, Amal Bouejla, Aldo Napoli, et Franck Guarnieri. 2012. « Integration of a bayesian network for response planning in a maritime piracy risk Management system ». In *Proceedings of 7th International Conference on System Of Systems Engineering*. Genoa, Italy.

Chemitte, Jérôme. 2008. « Adoption des technologies de l'information géographique et gestion des connaissances dans les organisations. Application à l'industrie de l'assurance pour la gestion des risques naturels ». École Nationale Supérieure des Mines de Paris.

Chen, Jin, Alan M. MacEachren, et Diansheng Guo. 2008. « Supporting the process of exploring and interpreting space-time multivariate patterns: the Visual Inquiry Toolkit ». *Cartography and Geographic Information Science* 35 (1): 33‑50.

Chi, Ed H. 2000. « A taxonomy of visualization techniques using the data state reference model ». In *Proceedings of the IEEE Symposium on Information Visualization*, 69‑76. Salt Lake City, USA: IEEE Computer Society.

Chuttur, Mohammad. 2009. « Overview of the Technology Acceptance Model: origins, developments and future directions ». *Sprouts: Working Papers on Information Systems* 9 (37).

Claramunt, Christophe, Thomas Devogele, Sébastien Fournier, Valérie Mallé-Noyon, Mathieu Petit, et Cyril Ray. 2007. « Maritime GIS: from monitoring to simulation systems ». *Information Fusion and Geographic Information Systems*, 34‑44.

Codd, EF, SB Codd, et CT Salley. 1993. *Providing OLAP (On-Line Analytical Processing) to User-analysts: An IT Mandate*. Vol. 32. Codd & Date, Inc.

Comfort, Louise K, Arjen Boin, et Chris C Demchak. 2010. *Designing Resilience: Preparing for Extreme Events*. Pittsburgh, Pa.: University of Pittsburgh Press.

Cottrell, Marie, et Patrick Letrémy. 2005. « How to use the Kohonen algorithm to simultaneously analyze individuals and modalities in a survey ». *Neurocomputing* 63 (janvier): 193‑207.

Cox, Louis Anthony Tony. 2008. « Some limitations of "Risk = Threat x Vulnerability x Consequence" for risk analysis of terrorist attacks ». *Risk Analysis* 28 (6): 1749‑61.

Daassi, Chaouki, Laurence Nigay, et Marie-Christine Fauvet. 2005. « A taxonomy of temporal data visualization techniques ». *Information-Interaction-Intelligence* 5 (2): 41‑63.

Dahlbom, Anders, et Tove Helldin. 2013. « Supporting threat evaluation through visual analytics ». In *2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, 155‑62.

Dalwadi, Neha, Bhaumik Nagar, et Ashwin Makwana. 2012. « Semantic Web and Comparative Analysis for Inference engines ». *International Journal of Computer Science and Information Technology* 3 (3): 3843‑47.

Davenport, Michael, et Chris Risley. 2006. *Information Visualization: the State of the Art for Maritime Domain Awareness*. DTIC Document.

Davies, John, Frank van Harmelen, et Dieter Fensel, éd. 2002. *Towards the Semantic Web: Ontology-driven Knowledge Management*. New York, NY, USA: John Wiley & Sons, Inc.

Davis, Fred D. 1993. « User acceptance of information technology: system characteristics, user perceptions and behavioral impacts ». *International Journal of Man-Machine Studies* 38 (3): 475‑87.

Davis, Fred D., Richard P. Bagozzi, et Paul R. Warshaw. 1989. « User acceptance of computer technology: a comparison of two theoretical models ». *Management Science* 35 (8): 982‑1003.

Demšar, Demšar, et Kirsi Virrantaus. 2010. « Space–time density of trajectories: exploring spatio-temporal patterns in movement data ». *International Journal of Geographical Information Science* 24 (10): 1527‑42.

Demšar, Urška. 2009. « Geovisualization and Geovisual Analytics ». In *The SAGE Handbook of Spatial Analysis*, édité par A. Stewart Fotheringham et Peter A. Rogerson, 41‑62. SAGE Publications Ltd.

Densham, Paul J. 1991. « Spatial Decision Support Systems ». In *Geographical Information Systems: Principles and Applications*, édité par David J. Maguire, Michael F. Goodchild, et David W. Rhind, 403‑12. London: Longman.

Devogele, Thomas. 2009. *Système d'information géographique temporelle maritime; Des distances linéaires à l'analyse temps réel des trajectoires*. Université de Bretagne Occidentale.

DiBiase, David, Alan M. MacEachren, John B. Krygier, et Catherine Reeves. 1992. « Animation and the role of map design in scientific visualization ». *Cartography and Geographic Information Science* 19 (4): 201‑14.

Directorate-General for Maritime Affairs and Fisheries. 2010. *Integrating Maritime Surveillance - Communication from the Commission to the Council and the European Parliament*. COM(2010) 584 final. European Commission. http://ec.europa.eu/maritimeaffairs/policy/integrated_maritime_surveillance/documents/integrating_maritime_surveillance_en.pdf.

Division on Technology and Logistics. 2013. *Review of Maritime Transport 2012*. Geneva, Switzerland: United Nations.

Doan, AnHai, Jayant Madhavan, Pedro Domingos, et Alon Halevy. 2002. « Learning to Map Between Ontologies on the Semantic Web ». In *Proceedings of the 11th International Conference on World Wide Web*, 662‑73. New York, NY, USA: ACM.

Dodge, Somayeh, Robert Weibel, et Ehsan Forootan. 2009. « Revealing the Physics of Movement: Comparing the Similarity of Movement Characteristics of Different Types of Moving Objects ». *Computers, Environment and Urban Systems* 33 (6): 419‑34.

Dörk, Marian, Sheelagh Carpendale, et Carey Williamson. 2012. « Visualizing explicit and implicit relations of complex information spaces ». *Information Visualization* 11 (1): 5‑21.

Dubée, Nicolas. 2001. « Sécurité des Systèmes d'Information : Analyse des risques ». septembre 19. http://download.secway.com/public/papers/presentations/analyse_risques/analyse_risques.pdf.

Dufour, Barbara, et Régis Pouillot. 2002. « Approche qualitative du risque ». *Épidémiologie et Santé Animale* 41: 35‑43.

Dumouchel, Anne Claire. 2013. « Gardes armés privés à bord des navires : la France change de cap ». *Bulletin Sentinelle*, décembre.

El Moussawi, Adnan. 2014. *Conception et développement d'un module de sélection de géovisualisations - Application à la sécurité maritime*. Beyrouth, Liban: ISSAE Cnam Liban.

Endsley, Mica R. 1995. « Toward a theory of situation awareness in dynamic systems ». *Human Factors* 37 (1): 32‑64.

Endsley, Mica R., Betty Bolte, et Debra G. Jones. 2003. *Designing for Situation Awareness: An Approach to User-Centered Design*. London ; New York: Taylor & Francis Ltd.

Engström, Johan, et Erik Hollnagel. 2007. « A General Conceptual Framework for Modelling Behavioral Effects of Driver Support Functions ». In *Modelling Driver Behaviour in Automotive Environments: Critical Issues in Driver Interactions with Intelligent Transport Systems*, édité par Carlo Cacciabue, 61‑84. Springer London Ltd.

Enguehard, René A., Rodolphe Devillers, et Orland Hoeber. 2013a. « Comparing interactive and automated mapping systems for supporting fisheries enforcement activities - a case study on vessel monitoring systems (VMS) ». *Journal of Coastal Conservation* 17 (1): 105‑19.

Enguehard, René A., Benjamin Fowler, Orland Hoeber, Rodolphe Devillers, et Wolfgang Banzhaf. 2012. « Integrating human knowledge within a hybrid clustering-classification scheme for detecting patterns within large movement data sets ». In *Proceedings of AGILE 2012*. Avignon, France.

Enguehard, René A., Orland Hoeber, et Rodolphe Devillers. 2013b. « Interactive exploration of movement data: A case study of geovisual analytics for fishing vessel analysis ». *Information Visualization* 12 (1): 65‑84.

Etienne, Laurent. 2011. « Motifs spatio-temporels de trajectoires d'objets mobiles, de l'extraction à la détection de comportements inhabituels - Application au trafic maritime ». Brest, France: Institut de Recherche de l'Ecole Navale.

Etienne, Laurent, Thomas Devogele, et Alain Bouju. 2012. « Spatio-Temporal Trajectory Analysis of Mobile Objects Following the Same Itinerary ». In *Advances in Geo-Spatial Information Science*, édité par Wenzhong Shi, Michael Goodchild, Brian Lees, et Yee Leung, 47‑57. CRC Press.

European Commission. 2008. *Integrated Maritime Policy for the EU - Working Document III on Maritime Surveillance Systems*. Ispra, Italy: Joint Research Centre. http://ec.europa.eu/maritimeaffairs/policy/integrated_maritime_surveillance/documents/maritime-surveillance_en.pdf.

Fabrikant, Sara Irina. 2001. « Building task-ontologies for geovisualization ». In . Beijing, China.

Fabrikant, Sara Irina, et Amy Lobben. 2009. « Introduction: Cognitive issues in geographic information visualization ». *Cartographica: The International Journal for Geographic Information and Geovisualization* 44 (3): 139‑43.

Fabrikant, Sara Irina, et André Skupin. 2005. « Cognitively Plausible Information Visualization ». In *Exploring Geovisualization*, édité par Jason Dykes, Alan M. MacEachren, et Menno-Jan Kraak, 667‑90. Elsevier Ltd.

Farmer, F. R. 1967. « Reactor safety and sitting: a proposed risk criterion »,. *Nuclear Safety* 8 (6): 539‑48.

Farrington, Jeanne. 2011. « Seven plus or minus two ». *Performance Improvement Quarterly* 23 (4): 113‑16.

Fayyad, Usama, Gregory Piatetsky-Shapiro, et Padhraic Smyth. 1996. « From Data Mining to Knowledge Discovery in Databases ». *AI magazine* 17 (3): 37.

Few, Stephen. 2004. « Tapping the Power of Visual Perception ». *Perceptual Edge*, septembre 4.

Fishbein, Martin, et Icek Ajzen. 1975. *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading, Mass: Addison-Wesley Pub.

Florez Guzman, Adrián Mauricio. 2012. *Simulation de la dérive d'objets flottants en mer*. Brest, France: Télécom Bretagne.

Fournier, Sébastien. 2005. « Intégration de la dimension spatiale au sein d'un modèle multi-agents à base de rôles pour la simulation: application à la navigation maritime ». Université de Rennes 1.

Foyle, David C., Anthony D. Andre, et Becky L. Hooey. 2005. « Situation awareness in an augmented reality cockpit: Design, viewpoints and cognitive glue ». In *Proceedings of the 11th International Conference on Human Computer Interaction*. Las Vegas, NV.

Frawley, William J., Gregory Piatetsky-Shapiro, et Christopher J. Matheus. 1992. « Knowledge discovery in databases: An overview ». *AI magazine* 13 (3): 57‑70.

Gallopoulos, Efstratios, Elias Houstis, et John R. Rice. 1994. « Computer as thinker/doer: problem-solving environments for computational science ». *IEEE Computational Science Engineering* 1 (2): 11‑23.

Gallopoulos, Stratis, Elias N. Houstis, et John R. Rice. 1992. *Future research directions in problem solving environments for computational science*. Washington, DC, USA.

Ganter, Bernard, et Rudolf Wille. 1999. *Formal Concept Analysis: Mathematical Foundations*. Berlin Heidelberg: Springer.

Gehlke, Charles E., et Katherine Biehl. 1934. « Certain effects of grouping upon the size of the correlation coefficient in census tract material ». *Journal of the American Statistical Association* 29 (185): 169‑70.

Glandrup, Maurice. 2013. « Improving Situation Awareness in the Maritime Domain ». In *Situation Awareness with Systems of Systems*, édité par Piërre van de Laar, Jan Tretmans, et Michael Borth, 21‑38. New York: Springer.

Goldsberry, Kirk, et Sarah Battersby. 2009. « Issues of change detection in animated choropleth maps ». *Cartographica: The International Journal for Geographic Information and Geovisualization* 44 (3): 201‑15.

Gómez-Pérez, Asunción, Mariano Fernandez-Lopez, et Oscar Corcho. 2004. *Ontological Engineering*. London: Springer-Verlag.

Goodchild, Michael F. 2010. « Twenty years of progress: GIScience in 2010 ». *Journal of Spatial Information Science*, n° 1: 3‑20.

Gotz, David, et Zhen Wen. 2009. « Behavior-driven visualization recommendation ». In *Proceedings of UI'09*, 315‑24. New York, USA: ACM.

Gotz, David, et Michelle X. Zhou. 2008. « Characterizing users' visual analytic activity for insight provenance ». In *IEEE Symposium on Visual Analytics Science and Technology, 2008. VAST '08*, 123‑30. Columbus, Ohio, USA.

Grau, Bernardo Cuenca, Christian Halaschek-Wiener, et Yevgeny Kazakov. 2007. « History matters: Incremental ontology reasoning using modules ». In *The Semantic Web, 6th International Semantic Web Conference*, 183‑96. Busan, Korea.

Grau, Bernardo Cuenca, Ian Horrocks, Boris Motik, Bijan Parsia, Peter Patel-Schneider, et Ulrike Sattler. 2008. « OWL 2: The next step for OWL ». *Web Semantics: Science, Services and Agents on the World Wide Web*, Semantic Web Challenge 2006/2007, 6 (4): 309‑22.

Gruber, Thomas R. 1993. « Toward principles for the design of ontologies used for knowledge sharing ». *International Journal Human-Computer Studies* 43: 907‑28.

Gruber, Thomas R. 2000. « What is an Ontology? ». http://www-ksl.stanford.edu/kst/what-is-an-ontology.html.

Guarino, Nicola. 1998. « Formal Ontology in Information Systems ». In *Proceedings of FOIS'98*, 3‑15. Trento, Italy: IOS Press.

Gudmundsson, Joachim, Patrick Laube, et Thomas Wolle. 2012. « Computational Movement Analysis ». In *Springer Handbook of Geographic Information*, édité par Wolfgang Kressp et David M. Danko, 725‑41. Berlin Heidelberg: Springer-Verlag.

Guo, Hanqi, Zuchao Wang, Bowen Yu, Huijing Zhao, et Xiaoru Yuan. 2011. « TripVista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection ». In *Proceedings of IEEE Pacific Visualization Symposium 2011*, 163‑70.

Haber, Robert B., et D. A. McNabb. 1990. « Visualization idioms: a conceptual model for scientific visualization systems ». In *Visualization in Scientific Computing*, édité par G. M. Nielson, B. Shriver, et L. J. Rosenblum, 74‑93. Los Alamitos, CA: IEEE Computer Society Press.

Han, Jiawei, et Micheline Kamber. 2006. *Data Mining: Concepts and Techniques*. 2nd éd. Morgan Kaufmann.

Harding, Jenny. 2013. « Usability of geographic information – Factors identified from qualitative analysis of task-focused user interviews ». *Applied Ergonomics* 44 (6): 940‑47.

Harnard, Stevan. 1994. « Computation is just interpretable symbol manipulation: cognition isn't ». *Special Issue on « What Is Computation » Minds and Machines* 4: 379‑90.

Harrower, Mark. 2007. « The cognitive limits of animated maps ». *Cartographica* 42 (4): 349‑57.

Harrower, Mark, et Cynthia A. Brewer. 2003. « ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps ». *The Cartographic Journal* 40 (1): 27‑37.

Hayward, Stan. 1984. *Computers for Animation*. Norwich UK: Page Bros. Ltd.

Healey, Christopher G., Kellogg S. Booth, et James T. Enns. 1996. « High-Speed Visual Estimation Using Preattentive Processing ». *ACM Transactions on Human Computer Interaction* 3 (2): 107‑35.

Heer, Jeffrey, et Maneesh Agrawala. 2006. « Software design patterns for Information Visualization ». *IEEE Transactions on Visualization and Computer Graphics* 12 (5): 853‑60.

Heer, Jeffrey, Michael Bostock, et Vadim Ogievetsky. 2010. « A tour through the visualization zoo. » *Communications of the ACM* 53 (6): 59‑67.

Heer, Thomas, Daniel Retkowitz, et Bodo Kraft. 2009. « Tool Support for the Integration of Light-Weight Ontologies ». In *Enterprise Information Systems*, édité par Joaquim Filipe et José Cordeiro, 175‑87. Springer.

Hernandez, Nathalie, et Josiane Mothe. 2006. *TtoO: une méthodologie de construction d'ontologie de domaine à partir d'un thésaurus et d'un corpus de référence*. IRIT/IRR.

Hilbert, Martin, et Priscila López. 2011. « The world's technological capacity to store, communicate, and compute information ». *Science* 332 (6025): 60‑65.

Hiltz, Starr R., et Murray Turoff. 1985. « Structuring computer-mediated communication systems to avoid information overload ». *Communications of the ACM* 28 (7): 680‑89.

Hofmann, Catherine, Hélène Richard, et Emmanuelle Vagnon. 2012. *L'âge d'Or des Cartes Marines : Quand l'Europe Découvrait le Monde*. Paris: Seuil.

Hollnagel, Erik. 1993. *Human Reliability Analysis: Context and Control*. London; San Diego, CA: Academic Press.

Hollnagel, Erik. 2000. « Modelling the orderliness of human action ». In *Cognitive engineering in the aviation domain*, édité par S. Sarter et R. Amalberti. Mahwah, NJ: Lawrence Erlbaum Associates.

Hollnagel, Erik. 2004. *Barriers and Accident Prevention*. Ashgate Publishing Limited.

Hollnagel, Erik, Jean Pariès, David Woods, et John Wreathall. 2010. *Resilience Engineering in Practice: A Guidebook*. Ashgate Publishing Limited.

Hollnagel, Erik, et David D. Woods. 1983. « Cognitive systems engineering: New wine in new bottles ». *International Journal of Man-Machine Studies* 18: 583‑600.

Hollnagel, Erik, et David D. Woods. 2005. *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering*. CRC Press.

Ho, Quan. 2013. « Architecture and Applications of a Geovisual Analytics Framework ». Linköping Universitet.

Ho, Quan, et Mikael Jern. 2013. « Geovisual Analytics and Storytelling Applied to a Flood Scenario ». In *GeoViz 2013*. Hamburg, Germany.

Horrocks, Ian, Peter Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, et Mike Dean. 2004. « SWRL: A Semantic Web Rule Language Combining OWL and RuleML ». http://www.w3.org/Submission/SWRL/.

Hurter, Christophe. 2010. « Caractérisation de visualisations et exploration interactive de grandes quantités de données multidimensionnelles ». Université Paul Sabatier - Toulouse III.

Hurter, Christophe, et Stéphane Conversy. 2007. « Extension d'un modèle de visualisation pour la caractérisation d'interfaces graphiques dynamiques ». In *Proceedings of the 19th International Conference of the Association Francophone d'Interaction Homme-Machine*, 39‑42.

Hurter, Christophe, Benjamin Tissoires, et Stéphane Conversy. 2009. « FromDaDy: Spreading aircraft trajectories across views to support iterative queries ». *IEEE Transactions on Visualization and Computer Graphics* 15 (6): 1017‑24.

ICA. 2014. « ICA Commission on GeoVisualization ». http://geoanalytics.net/ica/.

Idiri, Bilal. 2013. « Méthodologie d'extraction de connaissances spatio-temporelles par fouille de données pour l'analyse de comportements à risques - Application à la surveillance maritime ». Sophia Antipolis, France: Ecole Nationale Supérieure des Mines de Paris.

Idiri, Bilal, et Aldo Napoli. 2012a. « Découverte de règles d'association pour l'aide à la prévision des accidents maritimes ». In *Revue des Nouvelles Technologies de l'Information*, 23:243‑48. Bordeaux, France: Editions Hermann.

Idiri, Bilal, et Aldo Napoli. 2012b. « The automatic identification system of maritime risk accidents by rule-based reasoning ». In *Proceedings of 7th International Conference on System Of Systems Engineering*. Genoa, Italy.

IHO. 2000. *S-57 IHO Transfer Strandard for Digital Hydrographic Data*. International Hydrographic Organization.

IHO. 2010. *S-52 Specifications for Chart Content and Display Aspects of ECDIS*. International Hydrographic Organization.

IMO. 2013. *Amendments to the International Aeronautical and Maritime Search and Resacue (IAMSAR) Manual*.

IMO. 2014. « International Convention on Maritime Search and Rescue (SAR) ». *International Maritime Organization*. http://www.imo.org/About/Conventions/ListOfConventions/Pages/International-Convention-on-Maritime-Search-and-Rescue-(SAR).aspx.

Inselberg, Alfred. 1998. « Visual Data Mining with Parallel Coordinates ». *Computational Statistics* 13 (1): 47‑63.

International Atomic Energy Agency. 2014. « Potential risks, residual risks, acceptable risks ». http://www.iaea.org/ns/tutorials/regcontrol/assess/assess3212.htm.

International Standards Organisation. 1998. *ISO 9241-11:1998 - Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation (TEV) -- Partie 11: Lignes directrices relatives à l'utilisabilité*. Geneva.

International Standards Organisation. 2010. *ISO 9241-210:2010 Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems*. Geneva.

ISO. 2003. *ISO 19379:2003 - Ships and marine technology - ECS databases - Content, quality, updating and testing*. International Organization for Standardization.

ISO. 2009. *ISO Guide 73:2009 - Management du risque -- Vocabulaire*. International Organization for Standardization. http://www.iso.org/iso/fr/catalogue_detail?csnumber=44651.

Jackson, Peter. 1999. *Introduction to Expert Systems*. 3rd edition. Boston, MA, USA: Addison-Wesley.

Jankowski, Piotr, Timothy L. Nyerges, Alan Smith, T. J. Moore, et Emory Horvath. 1997. « Spatial group choice: a SDSS tool for collaborative spatial decisionmaking ». *International Journal of Geographical Information Science* 11 (6): 577‑602.

Journal Officiel. 1997. « Directive n° 96/82 du 09/12/96 concernant la maîtrise des dangers liés aux accidents majeurs impliquant des substances dangereuses ». janvier 14. http://www.ineris.fr/aida/consultation_document/1027.

Jung, Chin-Te, Chih-Hong Sun, et May Yuan. 2013. « An ontology-enabled framework for a geospatial problem-solving environment ». *Computers, Environment and Urban Systems* 38 (mars): 45‑57.

Kapler, Thomas, et William Wright. 2005. « GeoTime information visualization ». *Information Visualization* 4 (2): 136‑46.

Kéfi, Hajer, Michel Kalika, et Robert Reix. 2004. *Evaluation des Systèmes d'Information : une Perspective Organisationnelle*. Paris: Economica.

Keim, Daniel. 2002. « Information visualization and visual data mining ». *IEEE Transactions on Visualization and Computer Graphics* 8 (1): 1‑8.

Keim, Daniel, Jörn Kohlhammer, Geoffrey Ellis, et Florian Mansmann. 2010. *Mastering the Information Age - Solving Problems with Visual Analytics*. Goslar, Germany: Eurographics Association.

Keim, Daniel, Florian Mansmann, Jörn Schneidewind, et Hartmut Ziegler. 2006. « Challenges in visual data analysis ». In *10th International Conference on Information Visualization*, 9‑16. London, England.

Kergomard, Claude. 2009. « Notions de sémiologie (carto-)graphique ». novembre 30.

Kimerling, A. Jon, Aileen R. Buckley, Phillip C. Muehrcke, et Juliana O. Muehrcke. 2009. *Map Use: Reading and Analysis*. 6th edition. ESRI Press Academic.

Köhler, Wolfgang. 1964. *Psychologie de la Forme*. Paris: Gallimard.

Kosslyn, Stephen M. 1989. « Understanding charts and graphs ». *Applied Cognitive Psychology* 3: 185‑226.

Koua, Etien L., et Menno-Jan Kraak. 2005. « Evaluating Self-organizing Maps for Geovisualization ». In *Exploring Geovisualization*, édité par Jason Dykes, Alan M. MacEachren, et Menno-Jan Kraak, 627‑43. Elsevier Ltd.

Koussoulakou, Alexandra, et Menno-Jan Kraak. 1992. « Spatio-temporal maps and cartographic communication ». *The Cartographic Journal* 29 (2): 101‑8.

Kraak, Menno-Jan. 2003a. « Geovisualization illustrated ». *ISPRS Journal of Photogrammetry & Remote Sensing*, n° 57: 390‑99.

Kraak, Menno-Jan. 2003b. « The space-time cube revisited from a geovisualization perspective ». In *Proceedings of the 21st International Cartographic Conference*, 1988‑96. Durban, South Africa.

Kraak, Menno-Jan. 2006. « Beyond geovisualization ». *IEEE Computer Graphics and Applications* 26 (4): 6‑9.

Kraak, Menno-Jan. 2013. « GeoVisual Analytics and Cartography ». Keynote présenté à Séminaire Cartactive du GDR MAGIS, Paris, France, février 18.

Kraak, Menno-Jan, Rob Edsall, et Alan M. MacEachren. 1997. « Cartographic animation and legends for temporal maps: Exploration and or interaction ». In *Proceedings of the 18th International Cartographic Conference*. Vol. 1.

Kraak, Menno-Jan, et Ferjan Ormeling. 1996. *Cartography: Visualization of Geospatial Data*. Longman Group United Kingdom.

Kreuseler, Matthias, Norma López, et Heidrun Schumann. 2000. « A scalable framework for information visualization ». In *Proceedings of the IEEE Symposium on Information Vizualization*, 27‑36. Salt Lake City, Utah, USA.

Krygier, John. 1994. « Sound and Geographic Visualization ». In *Visualization in Modern Cartography (Modern Cartography Series)*, édité par Alan M. MacEachren et D. R. Fraser Taylor, 149‑66. New York: Pergamon.

Kveladze, Irma, et Menno-Jan Kraak. 2012. « What do we know about the space-time cube from cartographic and usability perspective? ». In *Proceedings of AutoCarto 2012*. Columbus, Ohio, USA.

Lagesse, Claire. 2011. *Conception et Développement d'un Géoportail Maritime*. Marne la Vallée, France: Ecole Nationale des Sciences Géographiques.

La Presse. 2014. « La piraterie «contenue» au large de la Somalie ». *La Presse*, janvier 23. http://www.lapresse.ca/international/afrique/201401/23/01-4731734-la-piraterie-contenue-au-large-de-la-somalie.php.

Laube, Patrick, Stephan Imfeld, et Robert Weibel. 2005a. « Discovering relative motion patterns in groups of moving point objects ». *International Journal of Geographical Information Science* 19 (6): 639‑68.

Laube, Patrick, Marc van Kreveld, et Stephan Imfeld. 2005b. « Finding REMO — Detecting Relative Motion Patterns in Geospatial Lifelines ». In *Developments in Spatial Data Handling*, 201‑15. Berlin Heidelberg: Springer.

Lavigne, Valérie, Denis Gouin, et Michael Davenport. 2011. « Visual analytics for maritime domain awareness ». In *Proceedings of IEEE International Conference on Technologies for Homeland Security*. Boston, USA.

Leclerc, Grégoire. 2002. *Poverty targeting with local and classic indicators: an attempt to integrate multiple perspectives*. Land Use Project. CIAT, Cali, Colombia.

Le Monde. 2013. « Paris va autoriser des gardes armés à bord des navires ». *Le Monde*. décembre 3. http://www.lemonde.fr/international/article/2013/12/03/paris-va-autoriser-des-gardes-armes-a-bord-des-navires_3524290_3210.html.

Liebig, Thomas, Christine Körner, et Michael May. 2009. « Fast visual trajectory analysis using spatial bayesian networks ». In *Proceeding of ICDM Workshops 2009, IEEE International Conference on Data Mining Workshops*, 668‑73. Miami, Florida, USA.

Lima, Manuel. 2013. *Cartographie des réseaux : L'art de représenter la complexité*. Paris: Eyrolles.

Liu, Shuangyan, Christopher Brewster, et Duncan Shaw. 2013. « Ontologies for crisis management: a review of state of the art in ontology design and usability ». In *Proceedings of the 10th Information Systems for Crisis Response and Management conference*, édité par T. Comes, F. Fiedrich, S. Fortier, J. Geldermann, et L. Yang. Baden-Baden, Germany.

Lloyd, David, Jason Dykes, et Robert Radburn. 2007. « Understanding geovisualization users and their requirements–a user-centred approach ». In *GIS Research UK 15th Annual Conference*, 209‑14. Maynooth, Ireland.

López-Pellicer, F. J., L. M. Vilches-Blázquez, J. Nogueras-Iso, Óscar Corcho, M. A. Bernabé, et A. F. Rodríguez. 2007. « Using a hybrid approach for the development of an ontology in the hydrographical domain ». In *2nd Workshop Ontologies for urban development: Conceptual models for practitioner*. Turin, Italy.

Lopez-Velasco, Céline, Jérôme Gensel, Marlène Villanova-Oliver, et Hervé Martin. 2009. « Vers une plate-forme de génération de SIG mobiles adaptés au contexte d'utilisation ». *Revue Internationale de Géomatique* 19 (2): 191‑210.

Lundblad, Patrik. 2013. « Applied Geovisual Analytics and Storytelling ». Linköping Universitet.

Luokkala, Pekka, et Kirsi Virrantaus. 2014. « Developing information systems to support situational awareness and interaction in time-pressuring crisis situations ». *Safety Science* 63 (mars): 191‑203.

Luqman, Faisal, et Martin Griss. 2010. « Overseer: a mobile context-aware collaboration and task management system for disaster response ». In *Proceedings of the 2010 8th International Conference on Creating, Connecting and Collaborating through Computing*, 76‑82. La Jolla CA, United States.

MacEachren, Alan M. 1995. *How Maps Work: Representation, Visualization, and Design*. 2nd revised edition. Guilford Press.

MacEachren, Alan M. 2001. *Cartography and GIS: Extending collaborative tools to support virtual teams*.

MacEachren, Alan M., et Isaac Brewer. 2004. « Developing a conceptual framework for visually-enabled geocollaboration ». *International Journal of Geographical Information Science* 18 (1): 1‑34.

MacEachren, Alan M., et David DiBiase. 1991. « Animated maps of aggregate data: Conceptual and practical problems ». *Cartography And Geographic Information Systems* 18 (4): 221‑29.

MacEachren, Alan M., Robert Edsall, Daniel Haug, Ryan Baxter, George Otto, Raymon Masters, Sven Fuhrmann, et Liujian Qian. 1999. « Virtual environments for geographic visualization: Potential and challenges ». In *8th ACM International Conference on Information and Knowledge Management*, 35‑40. ACM.

MacEachren, Alan M., Mark Gahegan, William Pike, Isaac Brewer, Guoray Cai, Eugene Lengerich, et F. Hardistry. 2004. « Geovisualization for knowledge construction and decision support ». *IEEE Computer Graphics and Applications* 24 (1): 13‑17.

MacEachren, Alan M., et Menno-Jan Kraak. 1997. « Exploratory cartographic visualization: Advancing the agenda ». *Computers & Geosciences* 23 (4): 335‑43.

Mackinlay, Jock. 1986. « Automating the design of graphical presentations of relational information ». *ACM Transactions on Graphics (TOG)* 5 (2): 110‑41.

Mackinlay, Jock, Pat Hanrahan, et Chris Stolte. 2007. « Show Me: automatic presentation for visual analysis ». *IEEE Transactions on Visualization and Computer Graphics* 13 (6): 1137‑44.

Maes, Pattie. 1994. « Agents that reduce work and information overload ». *Communications of the ACM* 37 (7): 30‑40.

Maggi, Fabrizio M., Arjan J. Mooij, et Wil M. P. van der Aalst. 2013. « Analyzing Vessel Behavior Using Process Mining ». In *Situation Awareness with Systems of Systems*, édité par Piërre van de Laar, Jan Tretmans, et Michael Borth, 133‑48. New York: Springer.

MAIB. 2014. *Report on the investigation of the collision between Paula C and Darya Gayatari in the south-west lane of the Dover Strait Traffic Separation Scheme on 11 December 2013*.

25/2014. United Kingdom: Marine Accident Investigation Branch. http://www.maib.gov.uk/cms_resources.cfm?file=/PaulaC_DaryaGayatri.pdf.

Malik, Abish, Ross Maciejewski, Yun Jang, Silvia Oliveros, Yang Yang, Ben Maule, Matthew White, et David S. Ebert. 2012. « A visual analytics process for maritime response, resource allocation and risk assessment ». *Information Visualization*, octobre.

Mallé-Noyon, Valérie. 2008. « Modèle de vue relative et spatio-temporelle de trajectoires géographiques d'objets mobiles : Application au contexte maritime ». *CFC*, décembre.

Mansouri, Mo, Roshanak Nilchiani, et Ali Mostashari. 2009a. « A risk management-based decision analysis framework for resilience in maritime infrastructure and transportation systems ». In *2009 3rd Annual IEEE Systems Conference*, 35‑41. Vancouver, Canada.

Mansouri, Mo, Brian Sauser, et John Boardman. 2009b. « Applications of systems thinking for resilience study in Maritime Transportation System of Systems ». In *2009 3rd Annual IEEE Systems Conference*, 211‑17. Vancouver, Canada.

McCollum, Barry. 2007. « A Perspective on Bridging the Gap between Theory and Practice in University Timetabling ». In *Practice and Theory of Automated Timetabling VI*, 23. Springer Berlin Heidelberg.

McNicholas, Michael. 2011. *Maritime Security: An Introduction*. Butterworth-Heinemann.

Meijerink, AMJ, CR Valenzuela, et A Stewart. 1988. *ILWIS: the integrated land and watershed management information system*. Scientific status report on the project GEO information system for land use zoning and watershed management. Enschede: ITC. http://library.wur.nl/WebQuery/groenekennis/278351.

Mendas, Abdelkader, Mohamed Amine Hamadouche, et F. Bouchenak. 2007. « Elaboration d'un système d'aide à la décision spatiale - Application à la dangerosité de l'infrastructure routière ». In *Colloque International VSST 2007*, 369‑91. Marrakech, Maroc.

Metz-Göckel, Hellmuth. 2012. « Gestalt Theory: History and Modernity ». In *Gestalt Theory and its Applications*. Helsinki, Finland. http://www.protsv.fi/lfs/luennot/2012_Metz-Gockel.pdf.

Miller, George A. 1956. « The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information ». *The Psychological Review* 63: 81‑97.

Mogford, Richard H. 1997. « Mental models and situation awareness in air traffic control ». *The International Journal of Aviation Psychology* 7 (4): 331‑41.

Monmonier, Mark. 1990. « Strategies for the visualization of geographic time-series data ». *Cartographica: The International Journal for Geographic Information and Geovisualization* 27 (1): 30‑45.

Morel, Michel. 2009. « SisMaris: Système d'Information et de Surveillance MARitime pour l'Identification des comportements Suspects de navires ». In *1ère Conférence Méditerranéenne Côtière et Maritime CM*, 2:261‑64.

Morel, Michel, et Stéphane Claisse. 2010. « Integrated system for interoperable sensors & information sources for common abnormal vessel behaviour detection & collaborative identification of threat (I2C) ». *IEEE Conference Publishing*.

Morel, Michel, Vincent Flori, Claire Saurel, Olivier Poirel, Aldo Napoli, Pierre Salom, et Gwenaele Proutière-Maulion. 2011. « TAMARIS: Traitement et Authentification des MenAces et RISques en mer ». In *Proceedings of Workshop Interdisciplinaire sur la Sécurité Globale*. Troyes, France.

Morel, Michel, Aldo Napoli, Jean-Pierre George, Florent Jangal, Marie-Annick Giraud, et Michel Botalla. 2010. « Surveillance et contrôle des activités des navires en mer : ScanMaris ». In *Proceedings of Workshop Interdisciplinaire sur la Sécurité Globale*. Troyes, France.

Moteff, John. 2005. *Risk Management and Critical Infrastructure Protection: Assessing, Integrating, and Managing Threats, Vulnerabilities and Consequences*. Order Code RL32561. CRS Report for Congress. Congressional Research Service.

Mountain, David. 2005. « Visualizing, Querying and Summarizing Individual Spatio-Temporal Behaviour ». In *Exploring Geovisualization*, édité par Jason Dykes, Alan M. MacEachren, et Menno-Jan Kraak, 181‑200. Elsevier Ltd.

Muehrcke, Phillip C. 1990. « Cartography and geographic information systems ». *Cartography And Geographic Information Systems* 17 (1): 7‑15.

Muller, Michel, Charles Fiterman, Daniel Tardy, Catherine Battut, et Jacques Dermagne. 2000. *Les causes et les conséquences du naufrage du pétrolier Erika*. Journal officiel de la République française, avis et rapports du Conseil économique et social.

Murakami, Yohei, Kazuhisa Minami, Tomoyuki Kawasoe, et Toru Ishida. 2002. « Multi-agent simulation for crisis management ». In *Proceedings of IEEE Workshop on Knowledge Media Networking, 2002*, 135‑39.

Napoli, Aldo. 2014. « Sécurité et sûreté de la maritimisation de l'énergie ». *MINES Revue des Ingénieurs*, avril.

Neerincx, Mark A., Stefan Kennedie, Marc Grootjen, et Franc Grootjen. 2009. « Modeling the Cognitive Task Load and Performance of Naval Operators ». In *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, édité par Dylan D. Schmorrow, Ivy V. Estabrooke, et Marc Grootjen, 260‑69. Lecture Notes in Computer Science 5638. Berlin Heidelberg: Springer-Verlag.

Nofi, Albert A. 2000. « Defining and measuring shared situational awareness », 74.

Noucher, Matthieu. 2009. « La donnée géographique aux frontières des organisations ». EPFL, Lausanne.

Noucher, Matthieu. 2013. « La donnée géographique aux frontières des organisations. Approche socio-cognitive de son appropriation. » *Cartes & Géomatique-Revue du Comité Français de Cartographie*, n° 215: 59‑62.

Noyes, Jan. 2001. « Human error ». In *People in Control: Human factors in control room design*, 3‑16. The Institution of Engineering and Technology, Michael Faraday House, Six Hills Way, Stevenage SG1 2AY, UK: IET.

Noyes, Jan, et Matthew Bransby, éd. 2001. *People in Control: Human Factors in Control Room Design*. The Institution of Engineering and Technology, Michael Faraday House, Six Hills Way, Stevenage SG1 2AY, UK: IET.

Observatoire du Monde Cybernétique. 2013. *Note 3ème trimestre 2013*. Délégation aux Affaires Stratégiques, Ministère de la Défense. http://www.defense.gouv.fr/content/download/222538/2476643/file/OMC2013T3.pdf.

Openshaw, Stan. 1983. *The Modifiable Areal Unit Problem*. Norwick: Geo Books.

Pala, Okan, et David Wilson. 2013. « User study analysis of a geovisualization decision support environment for critical infrastructure recovery ». In . Leuven, Belgium.

Parasuraman, Raja, et Victor Riley. 1997. « Humans and automation: use, misuse, disuse, abuse ». *Human Factors: The Journal of the Human Factors and Ergonomics Society* 39 (2): 230‑53.

Parasuraman, Raja, Thomas B. Sheridan, et Christopher D. Wickens. 2000. « A model for types and levels of human interaction with automation ». *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* 30 (3): 286‑97.

Parent, Christine, Stefano Spaccapietra, Chiara Renso, Gennady Andrienko, Natalia Andrienko, Vania Bogorny, Maria Luisa Damiani, et al. 2013. « Semantic Trajectories Modeling and Analysis ». *ACM Comput. Surv.* 45 (4): 42:1‑42:32.

Patel-Schneider, Peter, et Ian Horrocks. 2006. « OWL 1.1 Web Ontology Language Overview ». http://www.w3.org/Submission/owl11-overview/.

Patterson, Robert E., Leslie M. Blaha, Georges G. Grinstein, Kristen K. Liggett, David E. Kaveney, Kathleen C. Sheldon, Paul R. Havig, Timothy Lebo, et Jason A. Moore. 2014. « A human cognition framework for information visualization ». *Computers & Graphics* 42: 42‑58.

Peuquet, Donna J. 1994. « It's about time: a conceptual framework for the representation of temporal dynamics in geographic information systems ». *Annals of the Association of American Geographers* 84 (3): 441‑61.

Pfitzner, Darius, Vaughan Hobbs, et David Powers. 2003. « A unified taxonomic framework for information visualization ». In *Conferences in Research and Practice in Information Technology*, édité par J. Weckert, 1:57‑66. Canberra, Australia.

Power, Daniel J. 2002. *Decision Support Systems: Concepts and Resources for Managers*. Greenwood Publishing Group.

Pumain, Denise, et Thérèse Saint-Julien. 1997. *L'analyse spatiale : Localisations dans l'espace*. Cursus. Paris: Armand Colin.

Rasmussen, Jens. 1983. « Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models ». *IEEE Transactions on Systems, Man and Cybernetics* SMC-13 (3): 257‑66.

Ray, Cyril, Arnaud Grancher, Rémy Thibaud, et Laurent Etienne. 2013. « Spatio-temporal rule-based analysis of maritime traffic ». In *Conference on Ocean & Coastal Observation (OCOSS 2013)*. Nice, France.

Richard, Jean-François. 1982. « Mémoire et résolution de problèmes ». *Revue française de pédagogie*, 9‑17.

Riveiro, Maria. 2011. « Visual analytics for maritime anomaly detection ». University of Skövde.

Riveiro, Maria, et Göran Falkman. 2011. « The role of visualization and interaction in maritime anomaly detection ». In *Proceedings of SPIE-IS&T Electronic Imaging*, édité par P. Chung Wong, 7868:78680M 1‑12. San Francisco, CA, USA.

Riveiro, Maria, Göran Falkman, et Tom Ziemke. 2008. « Improving maritime anomaly detection and situation awareness through interactive visualization ». In *Proceedings of 11th International Conference on Information Fusion*, 1‑8.

Rivest, Sonia, Yvan Bédard, et Pierre Marchand. 2001. « Toward better support for spatial decision making: defining the characteristics of spatial on-line analytical processing (SOLAP) ». *Geomatica* 55 (4): 539‑55.

Roberts, Jonathan C., et Michael A. E. Wright. 2006. « Towards ubiquitous brushing for information visualization ». In *Proceedings of IEEE International Conference on Information Visualization*, 151‑56. London, England.

Robinson, Anthony C. 2008. « Design for Synthesis in Geovisualization ». The Pennsylvania State University.

Robinson, Anthony C. 2011. « Highlighting in geovisualization ». *Cartography and Geographic Information Science* 38 (4): 373‑83.

Roddick, John F., et Brian G. Lees. 2009. « Spatiotemporal Data Mining Paradigms and Methodologies ». In *Geographic Data Mining and Knowledge Discovery*, édité par Harvey J. Miller et Jiawei Han, 2nd éd., 33‑50. CRC Press.

Roth, Robert E. 2012. « Cartographic interaction primitives: framework and synthesis ». *The Cartographic Journal* 49 (4): 376‑95.

Royal Institution of Naval Architects. 2014. « Cruise safety dominates regulations following Concordia tragedy ». février. http://www.rina.org.uk/article1340.html.

Roy, Jean. 2008. « Anomaly detection in the maritime domain ». In *Optics and Photonics in Global Homeland Security IV*. Vol. 6945. SPIE.

Roy, Jean. 2009. « Automated reasoning for maritime anomaly detection ». In . La Spezia, Italy.

Roy, Jean, et Michael Davenport. 2009. « Categorization of maritime anomalies for notification and alerting purpose ». In . La Spezia, Italy.

Scheepens, Roeland, Huub van de Wetering, et Jarke J. van Wijk. 2014. « Non-overlapping Aggregated Multivariate Glyphs for Moving Objects ». In *2014 IEEE Pacific Visualization Symposium (PacificVis)*, 17‑24. Yokohama, Japan.

Schiewe, Jochen. 2013. « Geovisualisation and geovisual analytics ». *Kartographische Nachrichten*, n° Special Issue 2013 (juin): 122‑26.

Schneck, C. M. 2005. *Visual Perception. Occupational therapy for children*. Maryland Heigths: Mosby Elsevier.

Schreck, Tobias, Jürgen Bernard, Tatiana Tekusova, et Jörn Kohlhammer. 2009. « Visual cluster analysis of trajectory data with interactive Kohonen maps ». *Information Visualization* 8 (1): 14‑29.

Seppänen, Hannes, et Kirsi Virrantaus. 2010. « The role of GI-supported methods in crisis management ». *International Journal of Digital Earth* 3 (4): 340‑54.

Shneiderman, Ben. 1983. « Direct manipulation: A step beyond programming languages ». *Computer* 16 (8): 57‑69.

Shneiderman, Ben. 1996. « The eyes have it: a task by data type taxonomy for information visualizations ». In *Proceedings of the IEEE Symposium on Visual Languages*, 336‑43. Boulder, USA: IEEE Computer Society.

Simons, Daniel, et Christopher Chabris. 1999. *Selective Attention Test*. http://www.youtube.com/watch?v=vJG698U2Mvo&feature=youtube_gdata_player.

Skourup, Charlotte, et Arthur Aune. 2001. « Decision support in process control plants ». In *People in Control: Human factors in control room design*, 223‑37. The Institution of Engineering and Technology, Michael Faraday House, Six Hills Way, Stevenage SG1 2AY, UK: IET.

Slate. 2014. « Le nouvel épicentre de la piraterie mondiale se trouve de l'autre côté de l'Afrique ». *Slate.fr*. Consulté le août 22. http://www.slate.fr/story/83105/epicentre-piraterie-golfe-guinee.

Slocum, Terry A., Connie Blok, Bin Jiang, Alexandra Koussoulakou, Daniel R. Montello, Sven Fuhrmann, et Nicholas R. Hedley. 2001. « Cognitive and usability issues in geovisualization ». *Cartography and Geographic Information Science* 28 (1): 61‑75.

Snowden, David J., et Mary E. Boone. 2007. « A leader's framework for decision making ». *Harvard Business Review* 85 (11): 68‑76.

Sprague, Ralph H. 1980. « A framework for the development of decision support systems ». *MIS quarterly* 4 (4): 1‑26.

Staab, Steffen, et Rudi Studer. 2009. *Handbook on Ontologies*. 2nd éd. Berlin Heidelberg: Springer Verlag.

Stevens, S.S. 1946. « On the theory of scales of measurement ». *Science* 103 (2684): 677‑80.

Stolte, Chris, Diane Tang, et Pat Hanrahan. 2002. « Polaris: A system for query, analysis, and visualization of multidimensional relational databases ». *Visualization and Computer Graphics, IEEE Transactions on* 8 (1): 52‑65.

Strater, Laura D., Laurie A. Faulkner, John R. Hyatt, et Mica R. Endsley. 2006. « Supporting situation awareness under data overload in command and control visualizations ». In *International Ergonomics Association 2006 Congress*. Maastricht, Netherlands.

Studer, Rudi, V. Richard Benjamins, et Dieter Fensel. 1998. « Knowledge engineering: principles and methods ». *Data & Knowledge Engineering* 25: 161‑97.

Takatsuka, Masahiro, et Mark Gahegan. 2002. « GeoVISTA Studio: a codeless visual programming environment for geoscientific data analysis and visualization ». *Computers & Geosciences*, Shareware and freeware in the Geosciences II. A special issue in honour of John Butler, 28 (10): 1131‑44.

Thomas, James J., et Kristin A. Cook. 2005. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. National Visualization and Analytics Center.

Tomaszewski, Brian M., Anthony C. Robinson, Chris Weaver, Michael Stryker, et Alan M. MacEachren. 2007. « Geovisual analytics and crisis management ». In *Proceedings of the 4th International ISCRAM Conference*, édité par B. van de Walle, P. Burghardt, et Nieuwenhuis. Delft, Netherlands.

Tominski, Christian, C. T. H. Schumann, Martin Spindler, et Raimund Dachselt. 2011. « Towards utilizing novel interactive displays for information visualization ». In . Kobe, Japan.

Treisman, Anne, et Stephen Gormican. 1988. « Feature analysis in early vision: Evidence from search asymmetries ». *Psychological Review* 95 (1): 15‑48.

Tretmans, Jan, et Piërre van de Laar. 2013. « Introduction: Situation Awareness, Systems of Systems, and Maritime Safety and Security ». In *Situation Awareness with Systems of Systems*, édité par Piërre van de Laar, Jan Tretmans, et Michael Borth, 3‑20. New York: Springer.

Tufte, Edward R. 1983. *The Visual Display of Quantitative Information*. Cheshire, Conn.: Graphics Press USA.

Tufte, Edward R. 1997. *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, Conn.: Graphics Press USA.

Ullman, David G. 2007. « "OO-OO-OO!" The Sound of a Broken OODA Loop ». *The Journal of Defense Software Engineering*, avril, 22‑25.

Vandecasteele, Arnaud. 2012. « Modélisation ontologique des connaissances expertes pour l'analyse de comportements à risque - Application à la surveillance maritime ». Sophia Antipolis, France: Ecole Nationale Supérieure des Mines de Paris.

Vandecasteele, Arnaud, Rodolphe Devillers, et Aldo Napoli. 2014. « From movement data to objects behavior using semantic trajectory and semantic events ». *Marine Geodesy* 37 (2): 126‑44.

Vandecasteele, Arnaud, et Aldo Napoli. 2012. « Spatial ontologies for detecting abnormal maritime behaviour ». In *OCEANS 2012 MTS/IEEE Yeosu Conference: The Living Ocean and Coast - Diversity of Resources and Sustainable Activities*, 1‑7. Yeosu, South Korea.

Vandecasteele, Arnaud, et Aldo Napoli. 2013. « Modélisation ontologique pour l'analyse de comportements de navires à risques ». In *Colloque SAGEO 2013*. Brest, France.

Van der Vet, Paul E., et Nicolaas JI Mars. 1998. « Bottom-up construction of ontologies ». *IEEE Transactions on Knowledge and Data Engineering* 10 (4): 513‑26.

Van Hille, Pascal, Julie Jacques, Julien Taillard, Arnaud Rosier, David Delerue, Anita Burgun, et Olivier Dameron. 2012. « Comparing Drools and ontology reasoning approaches for telecardiology decision support ». In *Quality of Life through Quality of Information*, édité par J. Mantas, S. K. Andersen, M. C. Mazzoleni, B. Blobel, S. Quaglini, et A. Moen, 300‑304. IOS Press.

Van Laere, Joeri, et Maria Nilsson. 2009. « Evaluation of a workshop to capture knowledge from subject matter experts in maritime surveillance ». In *Proceedings of 12th International Conference on Information Fusion*, 171‑78.

Vatin, Gabriel, et Aldo Napoli. 2013a. « Guiding the controller in geovisual analytics to improve maritime surveillance ». In *Proceedings of 5th International Conference on Advanced Geographic Information Systems, Applications, and Services*, 26‑31. Nice, France.

Vatin, Gabriel, et Aldo Napoli. 2013b. « User profiling and geovisual analytics process modeling for maritime surveillance ». In *COST MOVE Workshop on Moving Objects at Sea*. Brest, France.

Vatin, Gabriel, et Aldo Napoli. 2013c. « Toward user-centred geovisual analytics in maritime surveillance ». In *GeoViz 2013*. Hamburg, Germany.

Vatin, Gabriel, et Aldo Napoli. 2013d. « High-level taxonomy of geovisual analytics tasks for maritime surveillance ». In *Proceedings of 26th International Cartographic Conference*. Dresden, Germany.

Vatin, Gabriel, et Aldo Napoli. 2013e. « Tâches d'exploration et d'analyse géovisuelle de données spatio-temporelles pour la gestion de risques ». In *Atelier Cartographie et Cognition - Conférence SAGEO 2013*.

Vatin, Gabriel, et Aldo Napoli. 2013f. « Formalisation des environnements d'analyse géovisuelle pour l'analyse de comportements à risques ». In *SCIGRAD'13*. Brest, France.

Vatin, Gabriel, et Aldo Napoli. 2014. « Using ontologies for proposing adequate geovisual analytics solutions in the analysis of trajectories ». In *18th International Conference Information Visualisation*. Paris, France: GraphicsLink.

Vatin, Gabriel, Eric Rigaud, et Aldo Napoli. 2014. « Geovisual analytics environment for supporting the resilience of maritime surveillance system ». In *4th International Engineering Systems Symposium*. Hoboken, NJ, USA.

Venkatesh, Viswanath, et Hillol Bala. 2008. « Technology Acceptance Model 3 and a research agenda on interventions ». *Decision Sciences* 39 (2): 273‑315.

Venkatesh, Viswanath, et Fred D. Davis. 2000. « A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies ». *Management Science* 46 (2): 186‑204.

Venkatesh, Viswanath, Michael G. Morris, Gordon B. Davis, et Fred D. Davis. 2003. « User acceptance of information technology: toward a unified view ». *MIS Quarterly* 27 (3): 425‑78.

Voiron, Christine, et Jean-Pierre Chéry. 2005. « Espace géographique, spatialisation et modélisation en Dynamique des Systèmes ». *Res-Systemica* 5.

Ware, Colin. 2000. *Information Visualization: Perception for Design*. Boston: Morgan Kaufmann.

Ware, Colin. 2008. *Visual Thinking for Design*. Morgan Kaufmann.

White, Roger, et Georgina Engelen. 2000. « High-resolution integrated modelling of the spatial dynamics of urban and regional systems ». *Computers, Environment and Urban Systems* 24 (5): 383‑400.

Wilkening, Jan, et Sara Irina Fabrikant. 2011. « How do decision time and realism affect map-based decision making? ». *Spatial Information Theory*, 1‑19.

Wilkinson, Leland. 1999. *The Grammar of Graphics*. New York: Springer Verlag.

Willems, Niels. 2011. « Visualization of vessel traffic ». Technische Universiteit Eindhoven.

Willems, Niels, Huub van de Wetering, et Jarke J. van Wijk. 2009. « Visualization of vessel movements ». In *Computer Graphics Forum*, 28:959‑66.

Willems, Niels, Huub van de Wetering, et Jarke J. van Wijk. 2011. « Evaluation of the visibility of vessel movement features in trajectory visualizations ». In *Computer Graphics Forum*, 30:801‑10.

Willems, Niels, Willem Robert van Hage, Gerben de Vries, Jeroen H.M. Janssens, et Véronique Malaisé. 2010. « An integrated approach for visual analysis of a multisource moving objects knowledge base ». *International Journal of Geographical Information Science* 24 (10): 1543‑58.

Wille, Rudolf. 1982. « Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts ». In *Ordered Sets*, édité par Ivan Rival, 445‑70. NATO Advanced Study Institutes Series 83. Springer Netherlands.

Wills, Graham, et Leland Wilkinson. 2010. « AutoVis: Automatic visualization ». *Information Visualization* 9 (1): 47‑69.

Woods, David D., Emilie M. Roth, et Kevin B. Bennett. 1990. « Explorations in joint human-machine cognitive systems ». In *Cognition, computing, and cooperation*, édité par Scott P. Robertson, Wayne W. Zachary, et John B. Black, 123‑58. Norwood, NJ, USA: Ablex Publishing Corp.

Wu, Anna, Xialong(Luke) Zhang, Gregorio Convertino, et John M. Carroll. 2009. « CIVIL: Support geo-collaboration with information visualization ». In *Proceedings of the ACM 2009 international conference on Supporting group work*, 273‑76.

Wybo, Jean-Luc. 2012. *Maîtrise des risques et prévention des crises*. Tec & Doc. Paris: Lavoisier.

Yan, Zhixian, Jose Macedo, Christine Parent, et Stefano Spaccapietra. 2008. « Trajectory Ontologies and Queries ». *Transactions in GIS* 12 (décembre): 75‑91.

Zhong, Chen, Chamseddine Zaki, Vincent Tourre, et Guillaume Moreau. 2010. « Event-based semantic visualization of trajectory data in urban city with a space-time cube ». In *Advances in Visualization, Imaging and Simulation*, édité par Nikos E Mastorakis et Valeri Mladenov, 99‑105. Faro, Portugal: WSEAS.

ANNEXES

1. Description des tâches modélisées

Les tableaux suivants présentent les concepts de tâches développés dans l'ontologie, ordonnés par groupes d'utilisation (quatre groupes parallèles dans l'étude du mouvement pour la gestion des risques). Dans chacun de ces tableaux, nous définissons le concept avec des termes plus simples, et nous donnons quelques exemples d'interaction ou de techniques l'utilisant.

1.1. Identification générale

Nom	Définition	Exemple(s)
T_DefineArea	Définition de la zone, par un ensemble de coordonnées ou un nom.	Choix d'un point / zone d'intérêt, zoom sur un cadre de coordonnées choisies, etc.
T_GetObjectProperties	Accès aux propriétés d'un objet sélectionné.	Fiche descriptive d'un objet
T_IdentifyObject	Identification d'un objet, par une propriété spécifique ou un identifiant.	
T_LocateObject	Localiser un objet spécifique, grâce à une propriété choisie.	Géolocaliser, choisir dans un tableau, etc.
T_RecognizePattern	Reconnaissance d'un motif connu, en comparant des données réelles à un modèle.	
T_RetrieveValue	Retrouver une valeur (numérique, qualitative) au sein d'un ensemble de données.	Champ de recherche, etc.
T_Sort	Trier des données alphanumériques.	Trier un tableau, une liste

1.2. Analyse spatio-temporelle

Nom	Définition	Exemple(s)
T_AnalyzePeriodicPattern	Analyser un motif périodique, afin de retrouver fréquence, durée, etc.	Vue circulaire (horloge)
T_CommonFlow	Retrouver une portion commune de mouvements dans un ensemble de trajectoires	
T_CompareDynamic	Comparer la dynamique (vitesse, rotation, etc.) de plusieurs objets, dans le temps.	Graphiques simultanés
T_GetDirectionChange	Retrouver les moments et lieux de changements de direction d'un objet mobile.	Graphique du cap dans le temps, taux de rotation, etc.
T_GetFrequentLocation	Retrouver les lieux régulièrement visités par des objets mobiles.	Densité de points
T_GetMeetingPlace	Retrouver un lieu de rencontre entre deux objets mobiles, ou plus.	Simulation de croisement, CPA, etc.
T_GetMeetingTime	Retrouver l'heure (instant) de rencontre entre deux objets mobiles, ou plus.	
T_GetOutliers	Retrouver les anomalies dans un jeu de données : mouvements, vitesse, attributs différents, etc.	Classification, visualisation simultanée de plusieurs objets et attributs, etc.
T_GetSpeedChange	Retrouver les moments et lieux de changements de vitesse d'un objet mobile.	Graphique de vitesse, d'accélération, etc.
T_GetStops	Retrouver les moments et lieux d'arrêts ou de ralentissements forts d'un objet mobile.	Cartographie des arrêts, graphe de vitesse, etc.
T_MeasureSpace	Mesure une distance ou une surface, dans l'espace 2D.	Règle et interaction cartographique
T_MeasureSpeed	Mesure la vitesse d'un objet mobile.	Graphe de vitesse
T_MeasureTimeDifference	Mesurer la différence de temps entre deux évènements.	Interaction avec un graphe
T_MeasureTimeDiffHeading	Mesurer la différence de temps entre deux caps identifiés pour un objet mobile.	
T_MeasureTimeDiffSpeed	Mesurer la différence de temps entre deux vitesses identifiés pour un objet mobile.	
T_PredictTrajectory	Déterminer la trajectoire à venir d'un objet mobile, à partir	Simulation de trajectoire, de dérive, etc.

	de son cap et sa vitesse courants.	
T_QuantifyDynamic	Donner un ordre de grandeur à la dynamique d'un objet mobile.	
T_QuantitySpatialRelation	Donner un ordre de grandeur à la proximité ou non de deux objets mobiles.	

1.3. Analyse du contexte

Nom	Définition	Exemple(s)
T_AnalyzeSpatialRelation	Analyser la proximité entre deux objets mobiles.	
T_CharacterizeAreaObjects	Caractériser les objets situés dans une zone définie	Attributs de vitesse, sens, attributs autres, etc.
T_CharacterizeAreaTraffic	Caractériser le trafic général dans une zone définie	Grands axes, sens, flux, vitesse générale, etc.
T_CharacterizeDistribution	Caractériser la répartition d'objets dans une zone définie	Densité, comptage,etc.
T_Cluster	Regrouper les objets proches dans l'espace.	
T_Generalize	Généraliser les objets géographiques.	Regroupement cartographiques de points,simplification de trajectoires
T_GetAreaInfo	Récupérer des informations sur la zone actuelle.	
T_GetMeteoMovement	Récupérer des informations météorologiques.	Vents, vitesses et hauteurs de vagues,etc.
T_SynthetizeTimeInfo	Synthétiser les informations temporelles.	Vue générale sur 24h pour un certain attribut, TimeLens,etc.
T_SynthetizeTrafficInfo	Synthétiser les informations géographiques de trafic.	Flux, origine-destination,etc.

1.4. Identification du risque

Nom	Définition	Exemple(s)
T_ExplainBehavior	Expliquer un comportement, par une investigation du mouvement.	
T_IdentifyRisk	Identifier les risques possibles, en comparant des scénarios.	
T_IdentifySuspects	Identifier des suspects, grâce aux objets proches ou croisés et au scénario identifié.	

2. Liste des visualisations formalisées

Le tableau suivant présente les noms des différents concepts développés dans l'ontologie, qui correspondents aux méthodes de visualisation pouvant être proposées par le système à base de connaissances. Ces méthodes ont été choisies dans la littérature scientifique et l'étude des outils de représentation et d'analyse d'objets mobiles. Aussi, certaines visualisations ont été choisies comme étant les plus simples (diagrammes, graphes, etc.), mais non nécessairement utilisées dans les systèmes actuels : celles-ci permettent donc de proposer des solutions simples à comprendre et à utiliser, mais qui ne sont pas encore développées dans les systèmes de surveillance ou d'analyse.

Nom	Définition	Exemples et références
VisSp_Choropleth	Cartes choroplèthes, aplats de couleurs	Technique traditionnelle de la cartographie, voir (Béguin et Pumain 1994)
VisSp_ClusterCircles	Regroupement de données ponctuelles par une représentation en cercles proportionnels	(Andrienko et Andrienko 2009 ; Andrienko et al. 2011 ; Koussoulakou et Kraak 1992)
VisSp_ClusterSquares	Regroupement de données ponctuelles par une représentation en carrés de couleur	idem
VisSp_CompositeDensityMap	Travaux par Willems, carte de densité par noyau utilisant deux rayons de recherche	Méthode présentée dans les publications (Willems et al. 2009 ; Willems et al. 2010 ; Willems 2011)
VisSp_CPA	*Closest Point of Approach*, méthode de détermination et représentation du point le plus proche sur deux trajectoires d'objets	
VisSp_DensityMap	Carte de densité	(Demšar et Virrantaus 2010 ; Willems 2011)
VisSp_FlowMap	Carte de flux de mouvements, selon un maillage défini : zones de départ, zones d'arrivée, effectif des flux	Technique traditionnelle de la cartographie (Béguin et Pumain 1994), avec certaines variantes (Andrienko et Andrienko 2013)
VisSp_Gantt	Diagramme GANTT pour la représentation d'activité ou de présence d'objets, dans le temps	
VisSp_HeadingGraph	Graphe de la direction (cap) d'un objet mobile, dans le temps	
VisSp_Histogram	Histogramme, diagramme présentant les effectifs de plusieurs classes de valeurs	
VisSp_KohonenMap	Carte de Kohonen, ou carte auto-adaptative, pour la	Méthode de statistiques et application à la cartographie

	classification d'objet à n dimensions	présentée dans plusieurs articles (Schreck et al. 2009 ; Cottrell et Letrémy 2005 ; Koua et Kraak 2005)
VisSp_ParallelCoordinates	Diagramme en coordonnées parallèles	(Andrienko et Andrienko 2001 ; Inselberg 1998 ; Parent et al. 2013)
VisSp_PieChart	Diagramme circulaire, en secteurs	
VisSp_RadarChart	Diagramme en radar, en étoile ou de Kiviat	Méthode utilisée dans les outils proposés dans (Riveiro 2011)
VisSp_RelativeTrajectory	Représentation de la trajectoire et vitesse d'un objet mobile, relativement à un autre objet mobile	Méthode présentée dans la publication (Mallé-Noyon 2008)
VisSp_Sankey	Diagramme de Sankey	
VisSp_ScatterPlot	Nuage de points, pour des valeurs numériques (X,Y)	
VisSp_ScatterPlotMatrix	Matrice de nuages de points, pour des données numérique multi-dimensionnelles	
VisSp_SmallMultiple	Petites cartes multiples, représentant un même attribut à plusieurs dates successives, ou plusieurs attributs à une même date	(Mackinlay et al. 2007 ; Goldsberry et Battersby 2009)
VisSp_SpaceTimeCube	Cube spatio-temporel	Méthode présentée dans les publications (Kraak 2003b ; Kveladze et Kraak 2012 ; Zhong et al. 2010)
VisSp_SpaceTimePrism	Prisme spatio-temporel, pour estimer l'espace éventuellement visité entre deux positions datées pour un objet mobile	Méthode dérivée du cube, et présentée dans la publication (Kraak 2003b)
VisSp_SpeedGraph	Graphe de la vitesse d'un objet mobile dans le temps	
VisSp_SpeedMap	Carte de vitesse représentant la trajectoire et la valeur de la vitesse d'un objet mobile	
VisSp_StopMap	Carte des arrêts et ralentissements d'un objet mobile, représentant la durée des stations	Méthode similaire présentée dans la publication (Buard et Brasebin 2012)
VisSp_TimeBar	Diagramme temporel en barres	
VisSp_TimeGraph	Graphe temporel linéaire	
VisSp_TimeLens	Lentille temporelle, méthode développée par Andrienko & Andrienko (2013)	Méthode présentée dans la publication (Andrienko et Andrienko 2013)
VisSp_TrajectoryMap	Carte des trajectoires d'un ou	

	plusieurs objet(s) mobile(s)	
VisSp_TrajectoryWall	Mur de trajectoires, méthode développée par Andrienko & Andrienko (2013)	Méthode présentée dans la publication (Andrienko et Andrienko 2013)
VisSp_WaveForceMap	Cartographie de la force des vagues	
VisSp_WindBarbsMap	Cartographie des vents par représentation barbules	
VisSp_WindSpeedMap	Cartographie de la vitesse des vents	

3. Modèle *Data State Reference Model* de Chi : exemples

Sur le tableau suivant, nous présentons la modélisation de Chi appliquée à certaines techniques de visualisation d'information, certaines spécifiques au domaine de l'information géographique. Ce tableau reprend chaque étape du modèle DSRM de Chi en colonnes, et les méthodes de visualisation en lignes. Ces exemples de modélisation ont été effectués en amont du développement de l'ontologie, afin d'identifier les étapes nécessaires et les mots-clés à utiliser lors de la mise au point du modèle formel par concepts, au sein de l'ontologie.

Méthode de visualisation	*Within Value*	*Data Transformation*	*Within Analytical Abstraction*	*Visualization Transformation*	*Within Visualization Abstraction*	*Visual Mapping Transformation*	*Within View*
Space-Time Cube	Trajectoires de n objets sur [T0, T1]	Extraire des données quantitatives	Abstraction : $(X,Y,Z)_n$; Filtre dynamique ; Normaliser	Représentation 3D des données quantitatives	>	Cube et carte géo, symboles linéaires	Rotation, Zoom, Animer, Changer les couleurs
Positions temps réel	Positions de n objets à l'instant t	Extraire des données quantitatives	Abstraction : $(X,Y)_n$; Filtre dynamique	Représentation 2D des données quantitatives	>	Carte géo, symboles ponctuels	Zoom, Déplacement, Changer les couleurs
Trajectoires temps réel	Positions datées de n objets sur [T0, T1]	Construire des données géométriques	Abstraction : lignes ; Filtre dynamique	Représentation 2D des lignes	>	Carte géo, symboles linéaires	Zoom, Déplacement, Changer les couleurs
Clustering positions	Positions de n objets à l'instant t	Regrouper les points proches (algo OPTICS)	Abstraction : groupements de points ; Changer paramètres de clustering	Contour des regroupements	Extraire l'enveloppe convexe	Carte géo, polygones	Zoom, Déplacement
Clustering trajectoires	Trajectoires de n objets sur [T0, T1]	Regrouper les trajectoires proches (algo OPTICS)	Abstraction : groupements de polylignes ; Changer paramètres de clustering	Contour des regroupements	Extraire l'enveloppe convexe	Carte géo, polygones	Zoom, Déplacement
Space-time prism	Trajectoires de n objets sur [T0, T1] U [T2, T3]	Extraire des données quantitatives ; Calculer les positions potentielles entre deux intervalles	Abstraction : $(X,Y,Z)_n$ + rayon/step/traj; Filtre dynamique ; Normaliser	Représentation 3D des données quantitatives	Créer le cube des positions potentielles	Cube et carte géo, symboles linéaires	Rotation, Zoom, Animer, Changer les couleurs

Heatmap	Positions de n objets	Extraire des données quantitatives	Abstraction : $(X,Y,Z)_n$; Filtre dynamique	Extraire la densité des points	Changer le rayon d'influence	Echelle de couleur thermique	Zoom, Navigation
Choroplète	Zones géométriques et données descriptives	Extraire des données quantitatives	Normaliser, agréger	Echelle de couleur pour les valeurs à représenter	>	>	Zoom, Changer l'échelle de couleurs
Carte de Kohonen	Données multidimensionnelles $(X1,\dots Xn)_m$	Réseau de neurones : dégager k vecteurs représentatifs	Classifier les vecteurs	Représentation dans un espace 2D	>	>	-

Formalisation d'un Environnement d'Aide à l'Analyse Géovisuelle Application à la Sécurité et Sûreté de la Maritimisation de l'Energie

RESUME : L'espace maritime est encore aujourd'hui le contexte de nombreux accidents et dangers, comme des collisions ou des attaques pirates. Afin de garantir le contrôle de la sûreté et de la sécurité de cet espace, il est nécessaire d'étudier les données de mouvement en temps réel (surveillance) et les évènements passés (analyse). Contrôleurs et analystes sont alors confrontés à de grandes quantités de données, qui doivent être étudiées grâce à des systèmes utilisant des cartes et autres visualisations. Cependant, ces outils sont limités en termes de capacités d'analyse. L'utilisation de méthodes d'analyse géovisuelle pourrait alors faciliter la reconnaissance de motifs, la détection d'anomalies et la découverte de connaissances. Toutefois, en raison de leur complexité d'utilisation, plusieurs de ces méthodes n'ont pas dépassé le stade académique, et ne sont pas encore utilisées de manière opérationnelle dans l'étude des risques maritimes.

Dans ce contexte, nous proposons un environnement d'aide à l'analyse géovisuelle, qui permet de guider l'utilisateur dans la visualisation et l'analyse d'informations pour l'étude des risques maritimes. Notre démarche de thèse se fonde sur la formalisation des cas d'utilisation, des utilisateurs et des méthodes de visualisation. Le recours à des ontologies et des règles permet de concevoir un système à base de connaissances, afin de proposer des méthodes adéquates pour la visualisation et l'analyse des trajectoires de navires. Nous illustrons cette proposition par plusieurs exemples d'analyse de risques en mer.

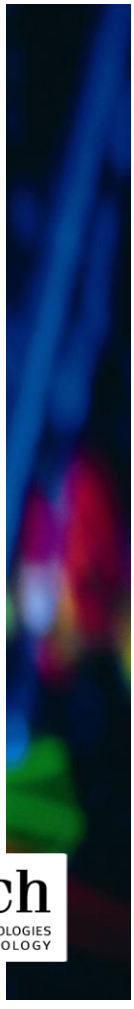
Mots clés : Analyse géovisuelle, géovisualisation, sécurité et sûreté, maritimisation de l'énergie, ontologies

Formalization of a Geovisual Analytics Support Environment Application to Maritime Oriented Energy Safety and Security

ABSTRACT: The maritime space is still a sensitive area due to many accidents and dangers, such as collisions or pirate attacks. In order to ensure the control of safety and security of this area, it is essential to study near real-time movement information (surveillance) or past events (analysis). Controllers and analysts are then faced to large sets of data, which must be studied with systems using maps and other visualizations. However, these tools are limited in terms of analysis capacities. Using geovisual analytics could be used to improve pattern identification, anomalies detection and knowledge discovery. However, due to the complexity of their use, most methods are still at the stage of research, and are not used yet in the operational word for studying maritime risks.

In this context, we propose a geovisual analytics support system to guide users in the visualization and the analysis of maritime risks. Our research methodology is based on the formalization of use cases, of users and of several visualization methods. Ontologies and rules are used to create a knowledge-based system, to select adequate solutions for the visualization and the analysis of ships' trajectories. Some examples for analyzing maritime risks are then presented to illustrate the use of such a system.

Keywords : Geovisual analytics, geovisualization, security and safety, maritime oriented energy, ontologies



MINES ParisTech

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY