

THÈSE

présentée

devant l'UNIVERSITE DE SAVOIE

pour l'obtention

du DIPLOME DE DOCTORAT

Discipline : Informatique

présentée et soutenue publiquement le 24/11/2006

par

CHRISTOPHE TRICOT

Cartographie Sémantique

Des connaissances à la carte



Directeur de thèse :

CHRISTOPHE ROCHE

JURY :

JEAN-PAUL BARTHES	Professeur des Universités	Président du jury
GUY MELANÇON	Professeur des Universités	Rapporteur
XAVIER BRIFFAULT	Chargé de recherche CNRS HDR	Rapporteur
ANNE DOURGNON-HANOUNE	Ingénieur recherche EDF R&D	Membre invité
ANDRE MANIFICAT	Directeur du GRETh INES-RDI	Membre invité
CHRISTOPHE ROCHE	Professeur des Universités	Directeur de thèse
DANIEL BEAUCHENE	Maître de Conférences	Codirecteur de thèse

Équipe Condillac « ingénierie des connaissances »

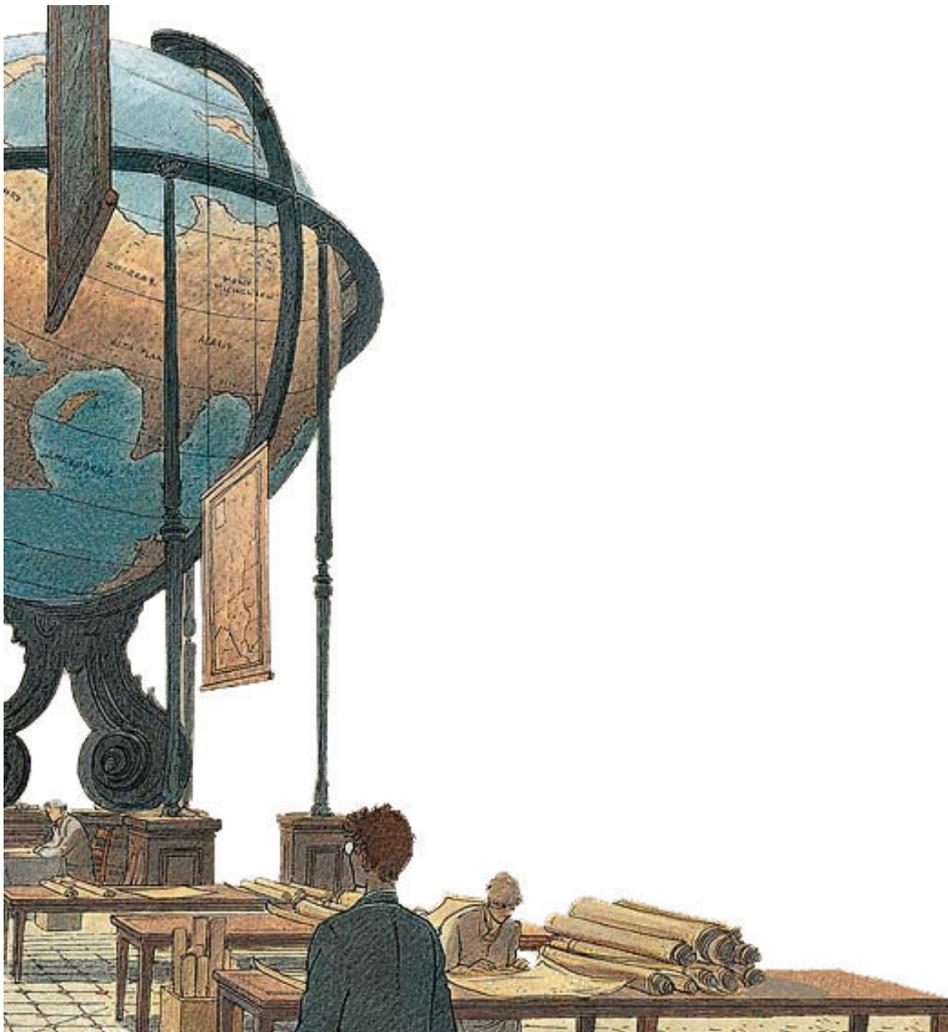
Laboratoire d'Informatique, Systèmes, Traitement de l'Information et de la Connaissance

Christophe Tricot



La Cartographie Sémantique

Des connaissances à la carte



À Nathalie...

Remerciements

Je souhaiterais adresser mes premiers remerciements à Christophe Roche, mon directeur de thèse, Professeur à l'Université de Savoie. Par son enthousiasme, ses compétences mais aussi la confiance qu'il m'a accordée, il m'a donné l'envie d'apprendre et de progresser. Durant ces années de collaboration, il a toujours su se montrer disponible et généreux. Ces quelques lignes sont bien peu de choses par rapport à tout ce qu'il m'a apporté. Je résumerais donc ma pensée en le citant : « C'est que du bonheur ! »

Merci aux personnes qui m'ont fait l'honneur d'accepter de participer à mon jury, Jean-Paul Barthès, Professeur des Universités comme président du jury, Guy Melançon, Professeur des Universités et Xavier Briffault, Chargé de recherche CNRS HDR comme rapporteurs, Anne Dourgnon Hanoune, Ingénieur recherche EDF R&D et André Manificat, Directeur du GRETh INES-RDI comme examinateurs, ainsi que Daniel Beauchêne, Maître de Conférences.

Merci à tous ceux qui ont contribué d'une manière ou d'une autre, à créer au sein de l'équipe Condillac un cadre de travail propice à l'étude et à la réalisation de cette thèse. Je remercie particulièrement Luc qui a su trouver les moyens avec « TeamI7 » pour me supporter et m'aider.

Merci à ma famille et à chacun de mes amis sans qui la vie n'aurait pas de sens. Ils ont toujours su m'encourager et me soutenir. Une pensée particulière pour mon « binôme » qui est à mes côtés depuis le début de cette aventure. J'espère que nos chemins ne cesseront pas d'être proches.

Avec une certaine émotion j'exprime mes plus profonds remerciements à mes parents qui ont su me donner toutes les chances pour réussir. Qu'ils trouvent dans la réalisation de ce travail l'aboutissement de leurs efforts ainsi que l'expression de ma reconnaissance.

Si je ne devais remercier qu'une personne ce serait sans hésitation Nathalie. Je ne la remercierais jamais assez pour son indéfectible soutien. Je lui dédie ce travail et je me réjouis chaque jour qu'elle ait acceptée de partager ma vie.

Sommaire :

RESUME ETENDU	1
PARTIE 1 : POSITIONNEMENT DE L'ETUDE	17
1.1 Contexte : la gestion des connaissances	18
1.2 Problématique : comment maîtriser un espace informationnel ?	19
1.3 Notre approche : cartographier pour maîtriser	20
1.4 Point de départ : besoins des utilisateurs	20
PARTIE 2 : ÉTAT DE L'ART	23
Rappel du plan	24
2.1 La cartographie	25
2.2 La cartographie de données abstraites	41
2.3 Synthèse & critiques	90
PARTIE 3 : PROPOSITIONS	91
Rappel du plan	92
3.1 La cartographie sémantique	93
3.2 Notre processus de cartographie sémantique	99
3.3 SNDF : notre formalisme de description pour la cartographie sémantique	109
3.4 La cartographie sémantique pour l'exploration	125
3.5 MDL : notre langage de cartographie sémantique	137
3.6 Notre architecture à agents et MVC distribué	147
3.7 Synthèse des propositions	156
PARTIE 4 : REALISATIONS DES PROPOSITIONS	157
Rappel du plan	158
4.1 Contexte	159
4.2 Cartographie Sémantique avec Os Map	163
4.3 Modélisation graphique avec SNCW	172
4.4 Exploration sémantique avec EyeTree et RadialTree	189
4.5 Synthèse des réalisations	195
PARTIE 5 : MISES EN ŒUVRE & VALIDATIONS EN GESTION DES CONNAISSANCES	197
Rappel du plan	198
5.1 Cartographie pour la gestion des compétences	199
5.2 Cartographie pour la navigation dans un fonds documentaire	210

5.3 Synthèse des mises en œuvres.....	220
PARTIE 6 : CONCLUSION & PERSPECTIVES.....	221
6.1 Synthèse et conclusion.....	222
6.2 Perspectives.....	223
PARTIE 7 : BIBLIOGRAPHIE.....	227
PARTIE 8 : ANNEXES	237
Rappel du plan	238
8.1 Lexique	239
8.2 Index.....	241
8.3 La perception visuelle	243
8.4 MDL.....	245
8.5 Description WSDL du service Os Map.....	254
8.6 Tables	256
8.7 Crédits images	260

Sommaire détaillé :

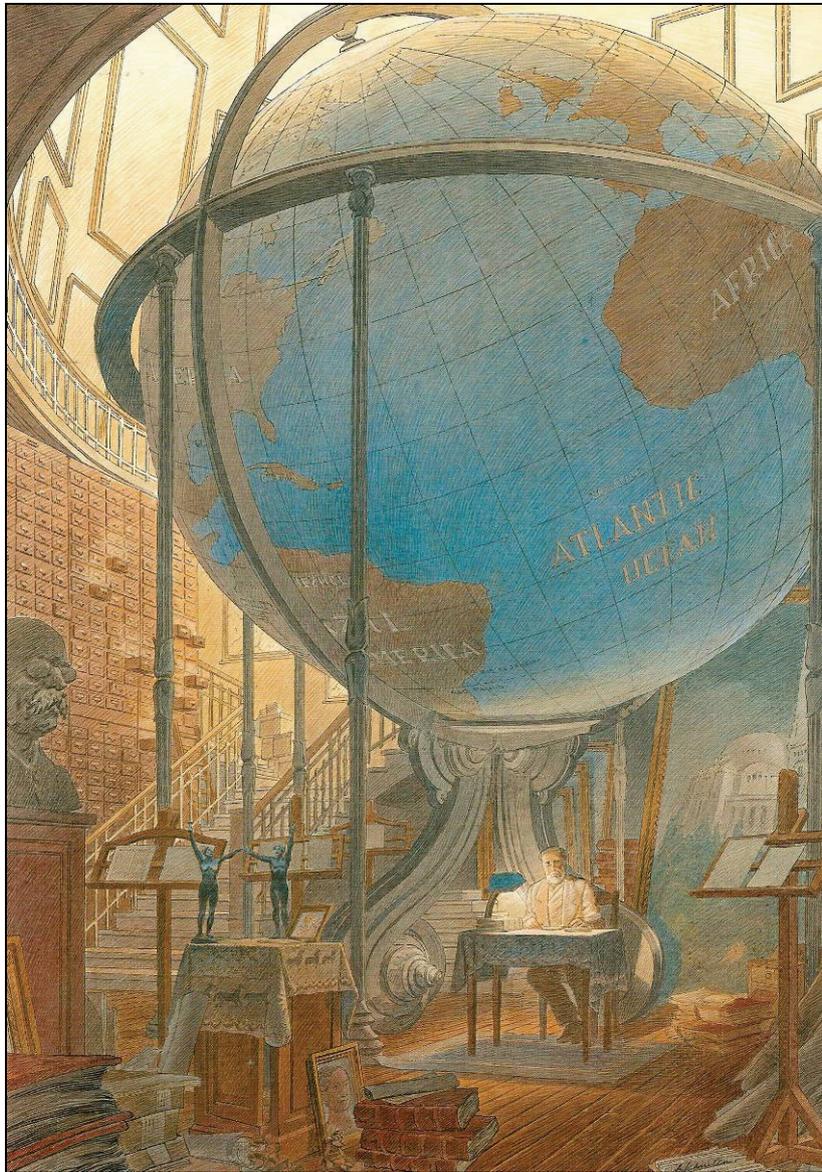
RESUME ETENDU	I
PARTIE I : POSITIONNEMENT DE L'ETUDE	17
1.1 Contexte : la gestion des connaissances	18
1.2 Problématique : comment maîtriser un espace informationnel ?	19
1.3 Notre approche : cartographier pour maîtriser	20
1.4 Point de départ : besoins des utilisateurs	20
PARTIE 2 : ÉTAT DE L'ART	23
Rappel du plan	24
2.1 La cartographie	25
2.1.1 Un monde de cartes	25
2.1.2 Quelques définitions	29
2.1.3 Principe de l'information cartographique.....	30
2.2 La cartographie de données abstraites	41
2.2.1 Une activité pluridisciplinaire	41
2.2.2 Les approches courantes	42
2.2.3 Un processus en niveaux	52
2.2.4 Les principaux paradigmes.....	54
2.3 Synthèse & critiques.....	90
PARTIE 3 : PROPOSITIONS.....	91
Rappel du plan	92
3.1 La cartographie sémantique	93
3.1.1 Cartographie orientée sémantique	94
3.1.2 Cartographie orientée retours d'expérience	97
3.1.3 Synthèse.....	98
3.2 Notre processus de cartographie sémantique	99
3.2.1 Des données à la carte	99
3.2.2 Des connaissances à la carte	101
3.2.3 Description du processus	103
3.2.4 Synthèse.....	107
3.3 SNDF : notre formalisme de description pour la cartographie sémantique.....	109
3.3.1 Nécessité d'un formalisme dédié	109
3.3.2 Les formalismes de description	110
3.3.3 Présentation de SNDF.....	116
3.3.4 Synthèse.....	123
3.4 La cartographie sémantique pour l'exploration	125
3.4.1 Positionnement de la problématique	125
3.4.2 Contexte et besoins.....	126
3.4.3 Cartographie de concepts et critères de cartographie	127
3.4.4 EyeTree	132
3.4.5 RadialTree	134

3.4.6 Synthèse.....	135
3.5 MDL : notre langage de cartographie sémantique	137
3.5.1 Nécessité d'un nouveau langage	137
3.5.2 Principe de MDL	138
3.5.3 La structure du langage.....	140
3.5.4 Exemples	144
3.5.5 Synthèse.....	146
3.6 Notre architecture à agents et MVC distribué	147
3.6.1 Cartographie sémantique orientée service.....	147
3.6.2 Distribution du processus de cartographie.....	148
3.6.3 Synthèse.....	154
3.7 Synthèse des propositions	156
PARTIE 4 : REALISATIONS DES PROPOSITIONS	157
Rappel du plan	158
4.1 Contexte.....	159
4.1.1 Équipe de Recherche Technologique	159
4.1.2 La démarche Ousia	160
4.2 Cartographie Sémantique avec Os Map.....	163
4.2.1 Objectif et principe	163
4.2.2 Services web et architectures orientées services	164
4.2.3 L'évocation du service	166
4.2.4 La mise en œuvre.....	167
4.2.5 Exemples	169
4.2.6 Synthèse.....	171
4.3 Modélisation graphique avec SNCW	172
4.3.1 La modélisation graphique de connaissances	173
4.3.2 Besoins pour la modélisation graphique	176
4.3.3 Le principe de SNCW	178
4.3.4 L'architecture de SNCW.....	182
4.3.5 Intégration dans KCW	188
4.3.6 Synthèse.....	188
4.4 Exploration sémantique avec EyeTree et RadialTree.....	189
4.4.1 Os EyeTree	189
4.4.2 Os RadialTree	192
4.4.3 Synthèse.....	194
4.5 Synthèse des réalisations	195
PARTIE 5 : MISES EN ŒUVRE & VALIDATIONS EN GESTION DES CONNAISSANCES	197
Rappel du plan	198
5.1 Cartographie pour la gestion des compétences.....	199
5.1.1 Problématique.....	199
5.1.2 Notre approche	200
5.1.3 Les cartes annuaires.....	202
5.1.4 Les cartes des emplois accessibles	203
5.1.5 Les cartes prévisionnelles	204
5.1.6 Les cartes stratégiques.....	206
5.1.7 Conclusion : la presse en parle.....	208

5.2 Cartographie pour la navigation dans un fonds documentaire.....	210
5.2.1 Problématique.....	210
5.2.2 Notre approche.....	212
5.2.3 Exploration graphique.....	212
5.2.4 Recherche graphique.....	214
5.2.5 Mise en œuvre : Misti.....	217
5.2.6 Conclusion.....	219
5.3 Synthèse des mises en œuvres.....	220
PARTIE 6 : CONCLUSION & PERSPECTIVES.....	221
6.1 Synthèse et conclusion.....	222
6.2 Perspectives.....	223
PARTIE 7 : BIBLIOGRAPHIE.....	227
PARTIE 8 : ANNEXES.....	237
Rappel du plan.....	238
8.1 Lexique.....	239
8.2 Index.....	241
8.3 La perception visuelle.....	243
8.4 MDL.....	245
8.4.1 DTD de MDL.....	245
8.4.2 Exemples de cartes avec MDL.....	248
8.5 Description WSDL du service Os Map.....	254
8.6 Tables.....	256
8.6.1 Table des figures.....	256
8.6.2 Table des tableaux.....	259
8.7 Crédits images.....	260
8.7.1 Les cités obscures.....	260
8.7.2 Circle Limite IV.....	260



Résumé étendu



« Si ardu étant ce chemin qu'il serait difficile de le donner à entendre par des paroles ou de l'écrire avec la plume. La meilleure explication qu'ait trouvé pour cela le génie des hommes est de le donner en peinture sur une carte. »

Martin Cortès

Cette introduction a pour objectif de résumer chacune des huit parties qui composent ce document, à savoir :

1. Positionnement de l'étude : dans un premier temps, nous situerons ce travail dans son contexte en présentant la problématique qui en est à l'origine. Ensuite nous définirons ce que nous entendons par la « cartographie sémantique », notre approche face à cette problématique.
2. État de l'art : pour présenter le domaine, nous exposerons un état de l'art critique et plus particulièrement les approches et les solutions existantes au regard de nos besoins.
3. Propositions : en réponse à cette problématique, nous décrivons l'ensemble de nos propositions qui loin d'être disparates constituent un tout cohérent c'est-à-dire, une méthode, un processus de cartographie sémantique, un formalisme de description, un langage de cartographie, deux nouveaux paradigmes pour l'exploration sémantique et enfin, une architecture d'agents logiciels.
4. Réalisations des propositions : la quatrième partie détaillera les principales réalisations de nos propositions.
5. Mises en œuvre & validations : puis, nous présenterons deux mises en œuvre qui ont permis de valider notre approche mais également de nous fournir de nombreux retours d'expériences sur l'application de la cartographie sémantique en milieu industriel.
6. Conclusions & perspectives : pour conclure, nous synthétiserons ce travail et nous identifierons quelques perspectives possibles.
7. Bibliographie : l'ensemble des travaux qui ont permis de réaliser ce travail sera listé dans cette partie.
8. Annexes : en fin de document, nous rassemblons les annexes constituées des différents éléments complémentaires à l'étude dont un lexique des termes importants.

1. Positionnement de l'étude

À l'origine de cette étude se trouve un constat important : les organisations (publiques ou privées) sont de plus en plus submergées par le nombre toujours croissant d'informations qu'il est nécessaire de maîtriser pour mener à bien leurs activités. En effet, étant obligées de faire face à la mondialisation des flux et des services, elles perçoivent de plus en plus que la maîtrise des informations constitue le nouvel enjeu à relever. Les organisations découvrent alors de nouveaux besoins et développent rapidement de nouvelles activités : veille stratégique, gestion de l'innovation, capitalisation des savoirs, gestion du capital intellectuel et humain, intelligence économique, etc. L'objectif est toujours le même : valoriser les informations. C'est l'objet même de la gestion des connaissances. Cette évolution marque l'entrée de nos sociétés dans l'ère de l'information et dans l'économie de la connaissance.

Problématique : comment maîtriser un espace informationnel ?

Maîtriser ses informations devient alors le challenge incontournable auquel doit faire face toute organisation indépendamment de sa taille. La conséquence directe est un accroissement permanent de la quantité d'information à manipuler.

Nous nommons « *espace informationnel* » l'ensemble des informations manipulées par une organisation. Pour chaque organisation, cet espace est très vaste. Il va des informations collectées (comme les informations sur les collaborateurs, les clients, les fournisseurs, les partenaires, les concurrents, les marchés, etc.) aux savoir-faire (comme les descriptions des outils et des procédés, la documentation scientifique et technique, les brevets, etc.).

Face à des espaces informationnels toujours plus importants, comment gérer toutes ces informations ? Comment faire face à autant de données ? Comment maîtriser un espace informationnel ? Ces préoccupations constituent la problématique à l'origine de ce travail.

Notre approche : cartographier pour maîtriser

La maîtrise d'un espace est une problématique universelle qui préoccupe l'homme depuis son origine. La cartographie résulte de cette problématique. Cartographier un espace permet de le visualiser en une carte pour mieux l'appréhender¹.

« Une image vaut mille mots ». Cette constatation sur la valeur de la visualisation est d'autant plus vraie si l'on travaille avec un grand nombre d'informations. La synthèse des travaux sur la cartographie (géographique) et sur la visualisation permet de considérer les cartes comme un médium idéal entre un grand nombre d'informations et l'esprit. Chaque carte joue alors le rôle de support externe de la pensée pour amplifier la cognition² de ses utilisateurs.

Face à ce constat, notre approche consiste à étudier les différents paradigmes de la cartographie et de les appliquer aux informations manipulées par les organisations pour construire des cartes. Notre objectif est alors d'aider les collaborateurs à appréhender leur espace informationnel pour mieux le maîtriser.

Nous nommons « cartographie de données abstraites », le recours aux paradigmes de cartographie pour représenter un espace informationnel. Bien que cette nomination soit discutable, nous la choisissons en référence à la « visualisation de données abstraites » définie par la communauté de la visualisation d'informations.

Point de départ : besoins des utilisateurs

Le nombre de paradigmes de cartographie est très important, par conséquent il est possible de construire un nombre quasiment infini de cartes différentes d'un même espace de données. Dans un tel contexte, comment construire une carte qui aidera effectivement ses utilisateurs à appréhender leur espace informationnel ?

Face à cette profusion de paradigmes, nous avons débuté par l'étude du résultat recherché. Ainsi, nous avons identifié plusieurs besoins que toute carte devrait satisfaire pour répondre à notre problématique d'appréhension d'un espace informationnel :

- > Naviguer selon la sémantique du domaine entendue comme l'ensemble des concepts permettant d'appréhender le domaine : pour leurs activités, les collaborateurs doivent naviguer dans l'espace informationnel de leur organisation (comme lors d'une recherche d'informations). Pour faciliter cette navigation, il est nécessaire que les cartes

¹ D'après le Trésor de la Langue Française, l'appréhension est la faculté de saisir par l'intelligence et plus particulièrement, toute opération intellectuelle relativement simple ou immédiate, soit de jugement, soit de mémoire, soit d'imagination, considérée comme s'appliquant à un contenu distinct de l'opération elle-même.

² Fonction complexe multiple regroupant l'ensemble des activités mentales (pensée, perception, action, volonté, mémorisation, rappel, apprentissage) impliquées dans la relation de l'être humain avec son environnement et qui lui permettent d'acquérir et de manipuler des connaissances (associations, rétroaction, traitement de l'information, résolution de problèmes, prise de décision etc.). (Le Grand Dictionnaire Terminologique)

véhiculent la sémantique du domaine. Si la cartographie respecte cette sémantique, les utilisateurs peuvent alors la comprendre, l'assimiler et l'exploiter. Utiliser une carte qui exploite la sémantique du domaine pour naviguer permet aux utilisateurs d'appréhender plus facilement l'espace informationnel.

- > Proposer une vision à plusieurs échelles : permettre aux utilisateurs de maîtriser un ensemble d'informations nécessite de leur fournir les moyens pour l'appréhender dans sa globalité mais aussi dans ses particularités. C'est pourquoi, une carte doit offrir simultanément une vision globale et synthétique de l'espace informationnel, et une vision particulière et détaillée de ce même espace.
- > Proposer une carte adaptée à l'utilisateur : tous les collaborateurs n'ont pas la même activité et le même niveau d'expertise. Chaque carte doit donc être adaptée à ses utilisateurs. De plus, une carte étant une représentation de l'espace informationnel, elle véhicule une signification. Il est alors nécessaire de bien choisir et adapter cette représentation pour maîtriser la signification qui sera finalement perçue.

Nous nommons « **cartographie sémantique** » la cartographie de l'espace informationnel d'une organisation répondant à ces besoins. Ainsi, elle devient une activité essentielle à la gestion des connaissances, permettant de tirer parti de toute la richesse des informations de l'organisation.

2. État de l'art

Au fil des siècles, la cartographie a perfectionné ses techniques et augmenté son champ d'application. À l'origine, la carte servait d'aide-mémoire en représentant graphiquement des éléments constitutifs d'un territoire limité (la cartographie antique). Ensuite, elle est devenue descriptive avec la couverture de l'ensemble du globe (la cartographie topographique). Puis, elle est devenue thématique en représentant des données de plus en plus « abstraites » (la cartographie thématique).

D'un point de vue général, la cartographie correspond à l'ensemble des études et des techniques qui permettent de construire une carte à partir d'informations. La carte est alors constituée de représentations graphiques de phénomènes (concrets ou abstraits) en positions relatives, localisables dans l'espace [Bertin, 1977; Bertin, 1999 (1^{re} éd. 1969)]. Cette opération de construction est menée par un cartographe. Il arrive que la carte soit utilisée comme méthode de recherche lorsqu'elle apporte elle-même une information nouvelle non encore prévue ou pressentie par son utilisateur. La carte devient alors un instrument de transmission de messages mais aussi un instrument de résolution [Poidevin, 1999].

La cartographie de données abstraites

Il existe de nombreux travaux qui traitent de la cartographie de données abstraites. Ils sont associés à deux communautés distinctes. La première, qui est aussi la plus ancienne, est la communauté de la « visualisation d'informations ». La deuxième est celle de la « visualisation de connaissances ». D'après leurs membres, ces deux communautés diffèrent par la nature des informations à cartographier. Pour la communauté de la visualisation d'informations, l'objectif est de visualiser graphiquement tous types de données abstraites. La cartographie est alors assimilée à un processus de transformation d'un ensemble d'informations en une carte. En revanche, pour la communauté de la visualisation de connaissances, l'objectif est de traiter uniquement des concepts. Ces derniers étant représentés par des données abstraites, leur cartographie exploite les techniques de la visualisation d'informations. Ainsi, la visualisation de connaissances peut être perçue comme une activité exploitant la visualisation d'informations à

des fins de gestion de connaissances. Quoi qu'il en soit, les deux communautés partagent les mêmes paradigmes.

La construction d'une carte à partir d'un ensemble d'informations n'est pas immédiate. En effet, il est tout d'abord nécessaire de représenter¹ les informations sous forme graphique avant de les visualiser². La cartographie d'un espace informationnel fait donc intervenir deux types de paradigmes : ceux qui permettent de représenter des informations (pour obtenir une représentation) et ceux qui permettent de visualiser ces représentations (pour obtenir une visualisation) sous forme de carte. Une fois la carte construite, l'utilisateur va interagir avec elle pour l'adapter à ses besoins. Toutes les techniques d'interactions constituent donc elles aussi un type de paradigme. Par conséquent, pour présenter les principaux paradigmes de la cartographie de données abstraites, nous avons choisi d'utiliser une classification principale basée sur le type des paradigmes : représentation, visualisation et interaction.

Synthèse & critiques

Notre objectif est d'aider les collaborateurs des organisations à appréhender leur espace informationnel au travers de cartes. D'après l'étude de l'existant, la cartographie de données abstraites permet de répondre en partie à cette problématique. En effet, il est déjà possible de construire une quasi-infinité de cartes pour un même espace. Cependant, nous avons besoin de cartes permettant aux utilisateurs de naviguer dans leur espace informationnel selon la sémantique de leur domaine avec une vision adaptée alliant différents niveaux d'échelles. Donc, pour répondre complètement à ces besoins, il subsiste alors deux problématiques :

- > Comment proposer une navigation basée sur la sémantique du domaine ? Les cartes doivent permettre aux utilisateurs de naviguer selon cette sémantique. Il est alors nécessaire qu'elle soit explicitée dans l'espace informationnel à travers une modélisation. La difficulté est alors de déterminer comment construire cette modélisation.
- > Comment déterminer les paradigmes à mettre en œuvre ? Les différentes approches ne nous indiquent pas comment fournir des cartes adaptées aux utilisateurs.

Les propositions associées à cette étude ont toutes pour objectif de répondre à cette double problématique.

3. Propositions

L'ensemble de nos propositions constitue un tout cohérent c'est-à-dire la cartographie sémantique proprement dite.

La première proposition est une **méthode** pour cartographier un espace informationnel selon nos besoins. Elle repose sur la prise en compte de la sémantique du domaine et sur les retours d'expérience. Nous nommons cette méthode la « cartographie sémantique ».

La seconde proposition est un **processus** de cartographie sémantique. La cartographie est alors assimilée à un ensemble d'opérations qui permettent de construire une carte d'un espace informationnel *brut* en passant successivement par un espace informationnel *structuré* puis un espace informationnel *représenté* et enfin un espace informationnel *visualisé*.

¹ Représenter signifie « rendre présent sous la forme d'un substitut, en recourant à un artifice » (Trésor de la Langue Française).

² Le sens de « visualiser » ne correspond pas à l'action de voir – percevoir par le sens de la vue - mais bien à « rendre sensible à la vue ce qui n'est pas naturellement visible » (Trésor de la Langue Française).

Ce processus de cartographie sémantique requiert un certain nombre de modélisations, doit répondre à des critères, et nécessite une architecture pour sa mise en œuvre.

C'est pourquoi, notre troisième proposition est notre **formalisme de description** dédié à la cartographie sémantique, nommé SNDF (Semantic Network Description Formalism). Ce formalisme permet de décrire d'une part les connaissances du domaine en un ensemble d'entités et de relations entre ces entités et d'autre part de décrire des propriétés de représentation pour ces entités et ces relations.

La quatrième proposition est la définition de **critères pour la cartographie** d'espaces informationnels *structurés* par une ontologie de domaine (comme l'utilisation de techniques « focus+context » avec une géométrie euclidienne) et la spécification de deux nouveaux paradigmes pour explorer un espace informationnel *structuré* : l'EyeTree et le RadialTree.

La cinquième proposition est un **langage** de haut niveau pour construire une carte en décrivant sa cartographie sémantique, nommé MDL (Map Description Language).

Enfin, la sixième et dernière proposition est une **architecture** de type système multiagents basée sur une extension du modèle MVC (Model View Controller). Elle permet d'intégrer la cartographie sémantique dans les solutions de gestion de connaissances à l'aide d'un ensemble d'agents logiciels.

Proposition 1 : la cartographie sémantique

Selon le principe de correspondance des représentations, une représentation graphique est d'autant plus efficace qu'elle correspond à la représentation mentale que se fait l'utilisateur de ces informations [Chabris & Kosslyn, 2005]. Pour un collaborateur, la représentation mentale d'un espace informationnel correspond à la sémantique de son domaine comprise comme l'ensemble des concepts permettant d'appréhender les objets du domaine. Pour être efficace, la carte doit donc respecter cette sémantique. Notre approche consiste à exploiter la sémantique du domaine pour déterminer les techniques à mettre en œuvre pour cartographier un espace informationnel.

Notre expérience de la cartographie nous a aussi révélé l'importance et la valeur des retours d'expérience pour construire une bonne carte. En effet, même si toutes les cartes sont basées sur la graphique (langage monosémique), une carte ne peut pas être considérée comme définitive et universelle. Depuis l'origine de la cartographie, les cartographes savent qu'une carte se doit d'être esthétique, intuitive et efficace [Tufte, 1986]. Malheureusement aucun système ne permet de garantir la satisfaction de ces trois propriétés et encore moins pour tous les individus. Face à cette problématique, il est donc important de toujours se laisser la possibilité de faire évoluer la cartographie selon les besoins des utilisateurs. Par conséquent, leurs retours d'expérience constituent un critère fiable pour juger de l'efficacité des cartes.

Nous proposons donc une cartographie basée d'une part, sur la sémantique du domaine pour déterminer les paradigmes à mettre en œuvre et d'autre part, sur la prise en compte des retours d'expérience pour améliorer et adapter les cartes. Nous qualifions notre approche de cartographie sémantique, c'est-à-dire la cartographie de l'espace informationnel d'une organisation basée sur sa sémantique.

Proposition 2 : notre processus de cartographie sémantique

La cartographie sémantique est un processus qui permet de passer d'un espace informationnel à une carte. En accord avec les approches courantes, l'espace informationnel que l'on souhaite cartographier doit tout d'abord être représenté (à l'aide d'opérations de représentation). On obtient alors un ensemble d'objets graphiques nommés « structures

visuelles ». Ensuite, cet « espace informationnel *représenté* » doit être visualisé (à l'aide d'opérations de visualisation) pour une obtenir une vue (la carte). La cartographie sémantique est donc assimilable à un ensemble d'opérations appliquées successivement à un espace informationnel que nous qualifions de « *brut* » (l'espace à cartographier). À chaque opération, l'espace informationnel est traduit vers un autre espace avec au minimum, un espace informationnel *représenté* et un espace informationnel *visualisé*.

Comme nous l'avons constaté, les approches courantes ne permettent pas de déterminer les techniques à mettre en œuvre pour obtenir une carte répondant à nos besoins. Le cœur de notre approche repose donc sur la prise en compte de la sémantique du domaine pour guider la cartographie selon les besoins des utilisateurs. Cet objectif implique d'explicitier la sémantique à l'aide d'une modélisation qui décrit les concepts du domaine ainsi que leurs relations.

Chaque collaborateur d'une organisation possède sa propre vision de son domaine et par conséquent, il possède aussi ses propres connaissances sur ce domaine. Il est donc nécessaire de proposer une cartographie capable de supporter plusieurs modélisations d'un même domaine correspondant à autant de points de vue théoriques différents. Pour illustrer la possibilité de concevoir plusieurs modélisations d'un même monde, il suffit de penser aux théories ondulatoires et corpusculaires de la lumière. Pour permettre aux utilisateurs de naviguer dans l'espace informationnel selon cette sémantique, il est alors nécessaire de restructurer l'espace informationnel de départ (*brut*) selon la modélisation du domaine. Le résultat est alors un nouvel espace que nous qualifions d'espace informationnel *structuré*.

Une fois structuré, l'espace informationnel doit être représenté graphiquement. C'est-à-dire qu'il est nécessaire de créer un nouvel espace constitué de structures visuelles. Comme nous le verrons dans notre état de l'art, il existe un très grand nombre de techniques de représentation pouvant être combinées pour construire une représentation. Chacune de ces représentations doit permettre de répondre à notre besoin d'adapter les cartes à l'utilisateur avec une navigation basée sur la sémantique du domaine. Chaque utilisateur possède un profil particulier (avec une activité et un niveau d'expertise donné). Une représentation unique ne peut pas convenir. Il est donc nécessaire de proposer un processus de cartographie capable de supporter plusieurs représentations d'un même espace informationnel correspondant à autant de cartes différentes. La modélisation permet alors de déterminer les opérations de représentations qui respectent la sémantique et par conséquent, elle permet de déterminer celles qui peuvent être utilisées.

Pour être perceptible par l'utilisateur, un espace informationnel *représenté* (par un ensemble de structures visuelles) doit être visualisé. Tout comme précédemment, cette opération de visualisation doit permettre de répondre à notre besoin d'adapter les cartes à l'utilisateur. En effet, chaque utilisateur possède un support particulier pour afficher la carte (écran, projecteur, papier, etc.) avec des dimensions tout aussi particulières. Une visualisation unique ne peut donc pas convenir. Il est alors nécessaire de proposer un processus de cartographie capable de supporter plusieurs visualisations d'un même espace informationnel *représenté* correspondant à autant de cartes différentes.

Lors de la navigation dans un espace informationnel, le point de vue des utilisateurs change tout comme leur activité et leur niveau d'expertise. De plus, il est possible qu'un même utilisateur appréhende simultanément un même espace informationnel selon plusieurs points de vue théoriques et pour différentes activités. Par conséquent, il est nécessaire que la cartographie ne soit pas un processus linéaire et unique. L'utilisateur doit pouvoir appréhender

son espace informationnel selon différents points de vue simultanément correspondants à autant de cartes différentes.

De plus, l'utilisateur doit pouvoir adapter ou changer les différentes opérations de chacune des cartes pour répondre à ses besoins. Dans ce dessein, le processus de cartographie doit mettre en œuvre des techniques d'interaction pour adapter les différents espaces. La cartographie d'un même espace informationnel selon différentes cartes adaptables par l'utilisateur permet de répondre à notre approche basée sur les retours d'expérience. À tout moment l'utilisateur doit avoir une carte adaptée.

En synthèse, la cartographie sémantique peut être assimilée à un ensemble d'opérations permettant de passer successivement par :

1. Un espace informationnel *brut* : il s'agit de l'espace initial que l'on souhaite cartographier.
2. Un espace informationnel *structuré* : cet espace est obtenu par des opérations de structuration de l'espace informationnel basée sur une modélisation construite au préalable.
3. Un espace informationnel *représenté* : cet espace est obtenu par des opérations de représentation de l'espace informationnel *structuré* (le choix des opérations est en accord avec la modélisation).
4. Un espace informationnel *visualisé* : cet espace est obtenu par des opérations de visualisation de l'espace informationnel *représenté* (le choix des opérations est, lui aussi, déterminé en accord avec la modélisation).
5. Une carte adaptée : l'utilisateur adapte ses cartes en interagissant avec les différents espaces de la cartographie.

Proposition 3 : SNDF, notre formalisme de description pour la cartographie sémantique

Le cœur de notre approche repose sur la prise en compte de la sémantique du domaine pour guider la cartographie. Cet objectif implique tout d'abord d'explicitier la sémantique à l'aide d'une modélisation puis de structurer l'espace informationnel *brut* selon cette modélisation (les concepts de la modélisation servent à décrire les éléments de l'espace *brut*). Tout ensemble d'informations (une modélisation ou un espace informationnel) se décrit selon une théorie. La difficulté est alors de définir une nouvelle théorie dédiée à la cartographie sémantique en accord avec notre processus et capable de répondre à nos besoins.

Pour cette problématique, nous proposons un formalisme de description, nommé SNDF (Semantic Network Description Formalism), permettant de décrire d'une part les connaissances du domaine en un ensemble d'entités et de relations entre ces entités et d'autre part de décrire des propriétés de représentation pour ces entités et ces relations.

Pour décrire les connaissances du domaine, SNDF a un double objectif. D'une part, il permet d'explicitier la sémantique du domaine en décrivant une modélisation et d'autre part, il décrit l'espace informationnel *structuré* selon cette modélisation. SNDF repose sur les principes des réseaux sémantiques et des systèmes à base de schémas. Des premiers, il exploite la capacité de décrire des connaissances par un ensemble d'entités et de relations entre ces entités et des seconds, il reprend la possibilité d'ajouter aux entités un ensemble d'attributs valués.

Pour permettre d'ajouter des propriétés de représentation pour chaque élément à cartographier, SNDF leur associe des informations dédiées à la cartographie comme une liste de descriptions de représentation (la description des futures structures visuelles) et une liste d'actions pour les interactions.

SNDF est notre formalisme de description dédié à la cartographie sémantique.

Proposition 4 : la cartographie sémantique pour l'exploration

La proposition quatre permet d'illustrer la cartographie d'un espace informationnel déjà structuré par une ontologie de domaine. En effet, la problématique de la navigation dans une base de connaissances a vivement motivé cette étude et plus particulièrement la navigation graphique dans une base de connaissances annotée sémantiquement par une ontologie de domaine. Par annotation sémantique, nous entendons l'indexation des éléments de la base de connaissances par les concepts de l'ontologie. Et par navigation graphique, nous entendons une exploration et une recherche au sein de la base de connaissances guidées par la modélisation du domaine sous la forme de visualisations interactives. Dans ce cas, la base de connaissances constitue un espace informationnel déjà structuré. La modélisation correspond à l'ontologie du domaine. La problématique est alors de définir la cartographie sémantique pour l'exploration d'un espace informationnel *structuré* selon une ontologie de domaine.

Pour résoudre cette problématique, nous avons appliqué une démarche en deux étapes. La première étape a été d'étudier et d'évaluer les différents paradigmes de visualisation existants. Pour cela, nous avons au préalable défini avec l'aide des utilisateurs les fonctionnalités attendues. La seconde étape a consisté quant à elle à réaliser plusieurs cartes interactives dont l'analyse des retours d'expérience nous a permis d'identifier de nouvelles fonctionnalités à mettre en œuvre.

Selon notre approche de cartographie sémantique, nous avons analysé les retours d'expérience des utilisateurs. Ils nous ont permis d'identifier, dans le cadre de notre application, quatre critères principaux pour la réalisation d'une cartographie sémantique interactive :

- > Utiliser une technique de type « focus + context » pour permettre à l'utilisateur de se concentrer sur certains éléments tout en facilitant l'accès aux autres éléments ;
- > Utiliser une géométrie euclidienne pour ne pas perturber la perception naturelle des manipulations du plan ;
- > Proposer une vue globale de l'ontologie (la modélisation) permettant à l'utilisateur de facilement appréhender l'ensemble de l'espace informationnel *structuré* ;
- > Pouvoir parcourir la base de connaissances tout en gardant un point fixe de référence.

Suite à ces retours d'expérience nous avons également pu identifier deux profils distincts d'utilisateurs : les « novices » et les « experts ». Les « novices » ont une certaine connaissance des concepts du domaine sans connaître exactement leur organisation ; alors que les « experts » ont une bonne maîtrise de l'ontologie de leur domaine. Ces deux profils n'ont pas les mêmes attentes concernant la vision globale de l'ontologie. Pour les novices, il est nécessaire de leur proposer une carte permettant une appréhension globale de cette modélisation. En revanche, il est nécessaire de proposer aux experts une carte de l'ontologie permettant de se situer rapidement dans l'espace informationnel. Ce constat nous a donc amené à réaliser deux cartographies différentes dédiées à la cartographie de bases de connaissances annotées sémantiquement : le « EyeTree » et le « RadialTree ».

EyeTree

Le « EyeTree » est une représentation avec déformation de type « Fisheye Polaire ». Il est basé sur une géométrie euclidienne offrant une vision globale de la structure conceptuelle. Cette technique fait partie des techniques de représentation avec déformation. Pour cela, les nœuds sont répartis radialement dans l'espace euclidien avant de subir une transformation via une fonction d'amplification continue appliquée aux coordonnées polaires des nœuds.

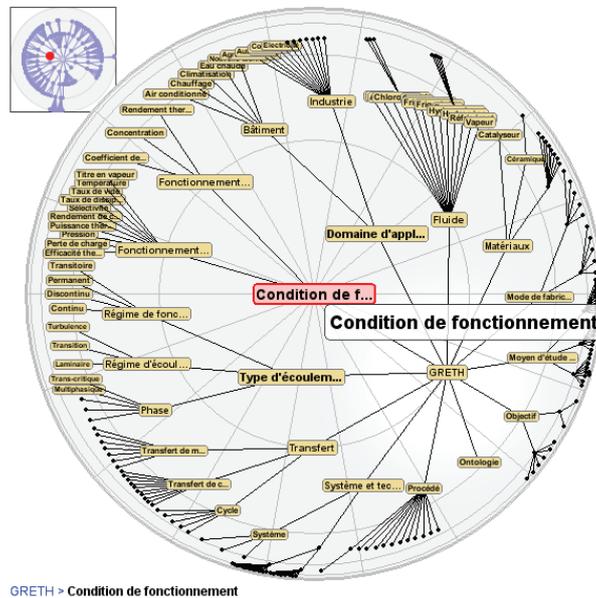


Figure 1 - Exemple de carte avec l'EyeTree.

Le résultat ressemble aux arbres hyperboliques [Lamping et al., 1995], mais les interactions de l'utilisateur (par exemple les translations) sont appliquées à un plan euclidien. Elles sont donc « naturellement » prévisibles par l'utilisateur. La transformation étant linéaire, le résultat n'est pas perturbant pour les utilisateurs (tout comme pour les « Perspective Wall » [Mackinlay et al., 1991]). La navigation s'effectue en parcourant la relation de généralisation/spécialisation entre les concepts et en glissant la projection de l'ontologie (« tree ») sur la représentation sphérique (« eye »), d'où le terme de « EyeTree ».

RadialTree

Par opposition à l'EyeTree, le « RadialTree » est plus particulièrement dédié aux experts du domaine. Il reprend les principales caractéristiques de l'EyeTree comme une cartographie globale de l'ontologie et l'utilisation d'une géométrie euclidienne. Cependant, le RadialTree n'applique pas de déformation globale à l'espace de représentation. En revanche, il propose une zone dite de « focus » permettant de magnifier une sous-partie de l'arbre.

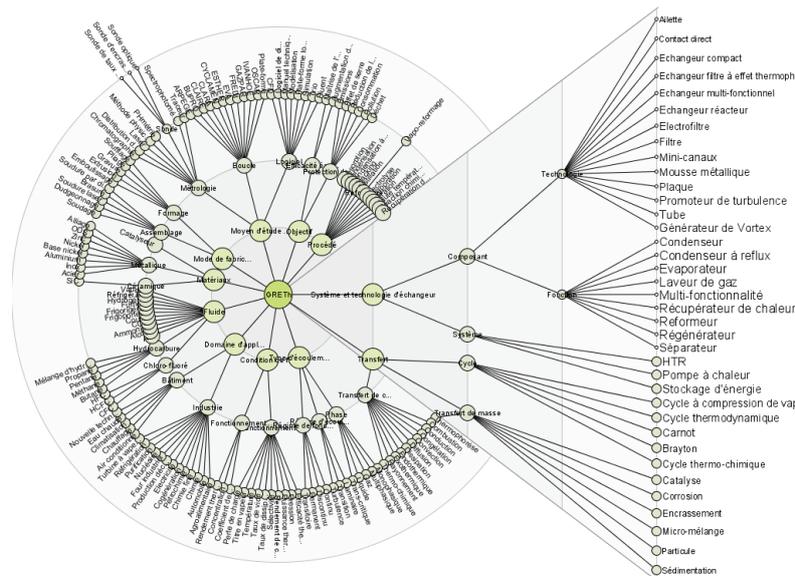


Figure 2 - Exemple de carte avec le RadialTree.

Les nœuds de l'arborescence sont répartis radialement dans l'espace euclidien. Les éléments situés dans la zone de focus sont magnifiés par une fonction d'amplification continue appliquée aux coordonnées. De plus, les feuilles de l'arbre situées dans la zone de focus sont projetées sur l'extrémité droite de cette zone et leur taille subit une déformation pour donner un effet équivalent à celui des menus fisheye augmentant ainsi leur lisibilité.

La réalisation de ces deux propositions est présentée avec les autres réalisations associées à cette étude.

Proposition 5 : MDL, notre langage de cartographie sémantique

La mise en œuvre de notre processus nécessite un langage pour décrire les cartes. Il existe de nombreux langages répondant à ce besoin. Il est possible de citer des langages comme Flash (langage propriétaire), SVG ou X3D (langages ouverts). Ces langages permettent de définir des cartes par la description d'un ensemble de structures visuelles. Par rapport à notre processus de cartographie sémantique, ces langages ne permettent de décrire que des espaces informationnels *représentés*. Ils ne permettent pas de décrire l'espace informationnel et encore moins de décrire l'espace *structuré* par une sémantique de domaine. Par conséquent, la définition de la carte n'est plus porteuse des connaissances cartographiées.

Pour répondre à notre problématique, nous avons donc été amenés à créer notre langage nommé MDL (pour « Map Description Language » ou « langage de description de cartes »). L'objectif de ce nouveau langage est de décrire une cartographie entendue comme l'ensemble des opérations nécessaires à appliquer à un espace informationnel *structuré* (en accord avec notre processus) pour construire une ou plusieurs cartes. L'utilisateur peut alors interagir avec les structures visuelles (comme les langages existants) mais aussi avec les connaissances (l'espace *structuré*). Pour décrire l'espace informationnel *structuré*, MDL exploite SNDF notre formalisme de description. D'un point de vue plus technique, MDL est un langage à balises de la famille XML.

Proposition 6 : notre architecture à agents et MVC distribué

Au fil des années, les systèmes d'information des organisations ont fortement muté. Ils sont passés pour la plupart de systèmes monolithiques à des systèmes de plus en plus répartis.

L'évolution majeure de ces dernières années dans les systèmes d'information est l'arrivée des services web qui permettent aux entreprises d'offrir leurs services comme des briques logicielles interconnectables. Chaque brique logicielle est assimilable à un agent logiciel évoluant sur Internet. Cette évolution permet avant tout de répondre à des besoins internes, mais elle permet aussi aux entreprises de sous-traiter certains de leurs traitements. La cartographie sémantique (entendue comme processus de gestion des connaissances) doit, elle aussi, être proposée comme un service pour intégrer les nouvelles architectures des systèmes d'information.

L'approche courante de cartographie de données abstraites consiste à appréhender la cartographie comme un processus séquentiel géré par une « boîte noire » permettant de passer d'un espace informationnel à une carte. Cette vision contraint à exécuter le processus complet de cartographie par une application sur le poste client et par conséquent, cela implique d'avoir un réseau performant pour transférer de grande quantité de données et des postes clients très performants pour supporter toute la cartographie.

Face à cette problématique, nous avons défini une nouvelle architecture dont le principe fondamental est la décomposition de la cartographie en différents services fournis par des agents logiciels. La finalité est de construire des cartes avec lesquelles les utilisateurs vont interagir. Pour permettre ces interactions, notre architecture met en œuvre des mécanismes d'interface homme-machine basés sur une extension du modèle MVC (Model View Controller) : notre « MVC distribué ». Cette évolution permet (à l'instar de l'original) de travailler simultanément sur plusieurs cartographies de sous-ensembles d'une même base constituant autant d'espaces informationnels cartographiés. Concrètement, nous avons décomposé notre architecture selon les principes de notre extension de MVC ce qui nous a permis de définir un ensemble d'agents logiciels qui interagissent entre eux pour construire des cartes.

4. Réalisations

Toutes les réalisations associées à cette étude ont été conçues dans une Équipe de Recherche Technologie (ERT) associée à une société de valorisation. « La recherche technologique a pour finalité de résoudre des problèmes posés par des entreprises industrielles ou de service ou liés à une demande sociétale. Elle se justifie par une demande du milieu socio-économique et ses thèmes sont donc inspirés par des objets industriels, produits, procédés, systèmes génériques ainsi que des services ». « Une équipe de recherche technologique (ERT) est une équipe de recherche qui, en partenariat avec des industriels, mène sur le moyen terme des recherches dans le cadre de projets visant à lever des verrous technologiques relatifs à des problèmes qui n'ont pas de solutions immédiates. Pour être reconnue, une ERT doit s'appuyer sur une recherche amont de qualité et justifier d'un engagement fort d'industriel(s) » (d'après les textes officiels¹). Par conséquent, une ERT est contrainte de réaliser d'une part, une recherche technologique pour lever des verrous technologiques et d'autre part, de valoriser ses travaux dans le cadre du partenariat industriel.

Les travaux associés à cette étude ont été réalisés au sein de l'équipe Condillac d'ingénierie des connaissances qui a été reconnue ERT par le Ministère de la Recherche. Le partenariat industriel s'effectue avec l'entreprise Ontologos Corp. qui s'est engagé à valoriser l'ensemble des travaux de recherche de l'équipe Condillac.

¹ <http://www.recherche.gouv.fr/technologie/rechped/default.htm>

La cartographie sémantique fait partie des verrous technologiques que l'équipe Condillac doit lever. Les réalisations présentées ici ont été guidées par la volonté de lever ce verrou, mais elles ont aussi été guidées par la volonté de répondre aux besoins du monde de l'entreprise. Les retours d'expérience de ces différentes applications constituent alors une validation efficace de notre approche.

Cartographie Sémantique avec Os Map

La première réalisation présentée est Os Map, notre service web de cartographie sémantique. Ce service met directement en œuvre l'ensemble de nos propositions. Tout d'abord, il exploite notre méthode et notre processus. Ensuite, il implémente une version de notre formalisme et de notre architecture à agents. Enfin, les cartes qu'il permet de construire sont décrites avec MDL, notre langage de cartographie et elles sont visualisées avec nos deux paradigmes d'exploration, l'EyeTree et le RadialTree.

Os Map est donc un service de cartographie sémantique conçu pour être intégré à toute application qui souhaite cartographier sémantiquement un espace informationnel.

Modélisation graphique avec SNCW

SNCW (Semantic Network Craft Workbench) correspond à notre deuxième réalisation. Il permet de modéliser graphiquement un domaine donné. Il exploite notre méthode de cartographie sémantique mais aussi notre formalisme de description et notre processus. Ainsi, il permet aux experts de modéliser leur domaine avec pour finalité la construction d'un système d'entités structurées capable d'exprimer leurs connaissances. Ce système produit alors ce que nous nommons une « base de connaissances ». SNCW est notre environnement graphique et interactif pour la modélisation et la représentation semi-formelle de connaissances.

SNCW permet de décrire plusieurs modélisations d'un même monde, mais aussi plusieurs vues d'une même modélisation (à chaque vue correspond une cartographie). La théorie que nous avons retenue est donc celle que nous proposons avec notre formalisme SNDF. Les bases de connaissances construites avec SNCW sont composées de modélisations et de cartographies de ces modélisations, les vues.

La réalisation de SNCW est basée sur un ensemble d'agents logiciels qui communiquent à l'aide de deux langages dédiés à la manipulation de bases de connaissances et de réseaux sémantiques. Les bases ainsi construites reposent sur un format ouvert de type XML, elles sont donc aisément partagées avec des applications tierces.

Exploration sémantique avec EyeTree et RadialTree

Enfin, les deux réalisations présentées sont l'EyeTree et le RadialTree correspondant à l'implémentation de notre spécification de paradigmes pour l'exploration sémantique d'espaces informationnels *structurés* ontologiquement.

Ces deux nouveaux paradigmes utilisent un langage de descriptions d'arborescences de type XML. Par conséquent, bien que ces deux paradigmes soient dédiés à la cartographie d'espaces informationnels structurés ontologiquement, ils permettent de cartographier les connaissances manipulées par des applications tierces pour cartographier tous types d'arborescences (comme SNCW).

5. Mises en œuvre & validations en gestion des connaissances

En accord avec sa fonction d'équipe de recherche technologique, l'équipe Condillac met en œuvre des solutions de gestion de connaissances pour lever différents verrous technologiques. L'appréhension d'un espace informationnel fait partie de ces verrous à lever et la cartographie sémantique est notre solution.

Pour témoigner de l'efficacité de notre approche, nous présentons deux cas concrets de sa mise en œuvre dans le monde industriel.

Le premier cas correspond à la cartographie sémantique appliquée à la gestion des compétences. Ce cas est à l'initiative de deux organisations : SADEC, un groupe d'expertise comptable et la ville de Cran-Gevrier. Cette étude sur la gestion des compétences a abouti à la réalisation d'OS-SKILL, environnement logiciel dédié à la gestion prévisionnelle des emplois et des compétences, étendues à la gestion des connaissances (GPECC). L'ensemble des fonctionnalités d'Os-Skill est intégré à un environnement de cartes interactives accessibles par Internet et intranet.

Le second cas correspond à l'utilisation de la cartographie sémantique pour gérer des fonds documentaires. Cette mise en œuvre est à l'initiative du GRETh, une émanation du CEA dédiée à la recherche sur les échangeurs thermiques.

Ces deux cas valident notre approche de cartographie sémantique en milieu industriel. Le déploiement et la mise en œuvre de la cartographie sémantique constituent une occasion formidable pour valider cette étude. Les retours des utilisateurs ainsi que ceux des commanditaires de ces mises en œuvre sont très positifs. Ils sont depuis convaincus que la cartographie sémantique constitue un atout pour leurs démarches de gestion de connaissances et des compétences.

6. Conclusion & Perspectives

La finalité de ce travail est de résoudre la problématique de l'appréhension par des collaborateurs de l'espace informationnel de leur organisation. Nous avons alors proposé la cartographie sémantique que nous définissons comme la cartographie de l'espace informationnel d'une organisation basée sur sa sémantique.

Sa mise en œuvre dans le monde des entreprises, au travers de la réalisation de nos propositions, a permis de valider notre approche. Chaque mise en œuvre nous permet de s'assurer que ce verrou technologique identifié par le Ministère de la Recherche est bien levé. Nous pouvons d'ailleurs nous réjouir d'avoir des témoignages sur l'apport bénéfique de notre approche dans la presse spécialisée. Fort de ces résultats, il a été estimé que le verrou de la cartographie sémantique a été levé et par conséquent que l'objectif de ce travail est atteint.

Néanmoins, nous identifions trois perspectives pour l'étude de la cartographie sémantique :

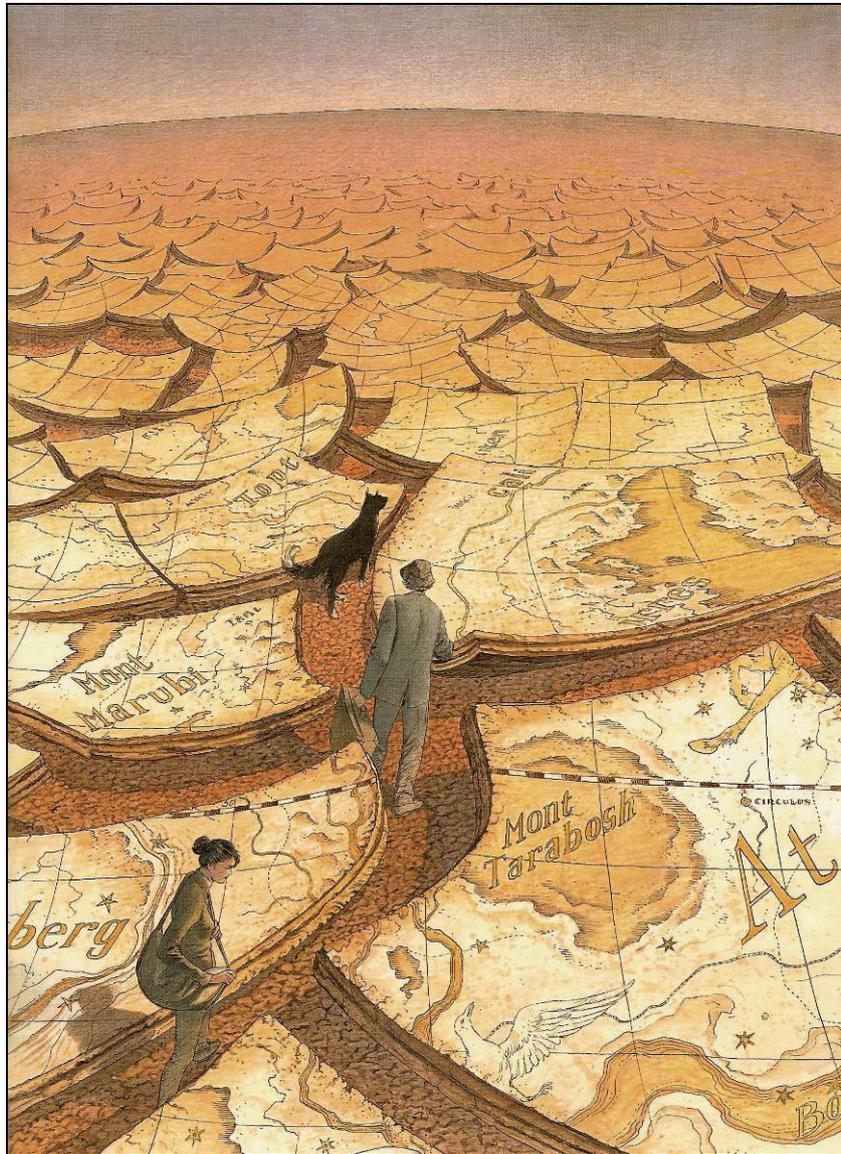
- > Évaluer les paradigmes : il ne faut pas omettre la dimension psychologique de la cartographie, c'est pourquoi nous estimons qu'il est nécessaire d'évaluer par des expérimentations l'impact de chaque paradigme sur les utilisateurs ainsi que leurs effets combinés. Ainsi en supposant une évaluation des paradigmes, de leur combinaison et de leurs impacts, leur choix pour constituer la cartographie peut alors être optimisé. Bien qu'une telle évaluation semble utopique dans l'état actuel de nos connaissances, nous estimons néanmoins que toute évaluation d'un paradigme permet de mieux en comprendre l'impact et par voie de conséquence de mieux l'utiliser.
- > Adapter les cartes à chaque utilisateur : les utilisateurs interagissent avec leurs cartes (comme déplacer le point de vue ou changer de zoom) pour naviguer dans l'espace

informationnel. Par soucis de construire des cartes adaptées, il nous paraît alors important d'étudier ces interactions. L'objectif final est de réduire au maximum le temps entre l'utilisation d'une carte et son adaptation c'est-à-dire adapter la cartographie en temps réel. Cette perspective de recherche est une approche de cognition située où le couple utilisateur/carte est à mettre en symbiose.

- > Définir de nouveaux paradigmes : pour augmenter les possibilités de cartographie et plus particulièrement pour changer les rapports entre l'utilisateur et l'espace informationnel il est nécessaire de définir de nouveaux paradigmes. Dans l'état actuel des choses, les cartes constituent une sorte d'interface entre les informations et l'utilisateur. Elle constitue alors le monde virtuel et l'utilisateur le monde physique. Pour être encore plus efficace, il est nécessaire de réduire au maximum la frontière entre ces deux mondes. Dans ce dessein nous identifions deux axes de recherche qui nous semblent très prometteurs : la « réalité virtuelle » et la « réalité augmentée ». La réalité virtuelle permet d'immerger l'utilisateur dans l'espace informationnel et la réalité augmentée permet d'immerger l'espace informationnel dans le monde réel.

Nous pensons qu'à l'avenir la recherche dans ces trois perspectives offriront de nouveaux travaux pour construire des cartes permettant aux utilisateurs de toujours mieux maîtriser leurs espaces informationnels ; c'est l'objectif de la cartographie sémantique.

Partie I : **Positionnement de l'étude**



« [Ce livre] aborde une problématique à laquelle sont confrontés quotidiennement beaucoup de professionnels : celle de comprendre et de gérer leurs territoires grâce à la cartographie. L'enjeu est grand : meilleurs seront les documents cartographiques, plus sûres seront les décisions qui en émanent. »

Didier Poidevin

Que ce soit de simples parcelles de terre, un pays ou un continent, l'homme a toujours eu besoin de gérer des territoires. Pour l'aider, il a inventé la cartographie. Elle lui permet de construire des cartes et ainsi d'appréhender plus facilement chaque territoire.

De nos jours, les organisations rencontrent la même problématique, mais avec des territoires de nature différente. Elles doivent gérer leur patrimoine intellectuel c'est-à-dire leurs connaissances explicites (les documents, projets, retours d'expériences, etc.) et leurs connaissances tacites (les expériences, le savoir-faire, etc.). Pour répondre à cette problématique de gestion de territoires, ce travail propose une nouvelle cartographie appliquée aux organisations : la cartographie sémantique.

I.1 Contexte : la gestion des connaissances

À l'origine de cette étude se trouve un constat important : les organisations (publiques ou privées) sont de plus en plus submergées par le nombre toujours croissant d'informations qu'il est nécessaire de maîtriser pour mener à bien leurs activités. En effet, étant obligées de faire face à la mondialisation des flux et des services, elles perçoivent de plus en plus la maîtrise des informations comme le nouvel enjeu à relever. Cette évolution marque l'entrée de nos sociétés dans l'ère de l'information.

La fin de l'ère industrielle et le début de l'ère de l'information. L'ère industrielle a été celle de l'évolution des outils. Les organisations et plus particulièrement les entreprises rivalisaient sur le plan technique pour prospérer. Elles se livraient une guerre pour perfectionner leurs outils de production et les innovations étaient majoritairement techniques et productives. La compétitivité, la valeur et la puissance des entreprises étaient alors directement fonctions de celles des outils et moyens de production. Les informations n'avaient alors pas beaucoup de valeur reconnue.

Puis, les entreprises ont spécialisé leurs activités et les marchés ont fusionné en un marché mondial unique. Il leur est devenu de plus en plus difficile d'être compétitif et de se démarquer des concurrents. C'est dans ce contexte mondialisé, que la prise de conscience de la valeur stratégique des informations a émergé. En effet, à moyen de production égal et sous réserve de maîtriser ses informations, l'organisation qui possède les meilleures connaissances (comme le savoir-faire) est stratégiquement la plus puissante. La compétitivité se déplace alors du plan industriel sur le plan informationnel. Cette évolution marque le passage de l'ère industrielle à l'ère de l'information.

La gestion des connaissances. Les entreprises se trouvent alors face à une nouvelle problématique : comment gérer les informations nécessaires à l'activité ? Comment les valoriser ? Comment déterminer les informations les plus pertinentes et extraire de nouvelles connaissances ? L'informatique, qui a prouvé son potentiel durant la deuxième guerre mondiale et pendant la « guerre froide », est mise à contribution et prend alors l'essor qu'on lui connaît. Simultanément, les entreprises découvrent de nouveaux besoins et développent rapidement de nouvelles activités : veille stratégique, gestion de l'innovation, capitalisation des savoirs, gestion du capital intellectuel et humain, intelligence économique, etc. La conséquence directe est un accroissement permanent de la quantité d'information à manipuler.

Toutes ces activités sont regroupées sous le nom de *gestion des connaissances*. Elle consiste à capturer et à représenter les connaissances¹ des organisations pour faciliter leur accès, leur partage et leur réutilisation [Dieng-Kuntz et al., 2005]. D'après Grundstein Michel (cité dans [Dieng-Kuntz et al., 2005]), la finalité est :

« localiser et rendre visible les connaissances de l'entreprise, être capable de les conserver, y accéder et les actualiser, savoir comment les diffuser et mieux les utiliser, les mettre en synergie et les valoriser ».

De nombreux auteurs ont proposé leur méthode pour gérer les connaissances (liste non exhaustive de ces méthodes dans [Dieng-Kuntz et al., 2005]). Dans la plupart des cas, elles alternent des phases de création, de stockage, de transfert et d'exploitation d'informations. À chaque fois il est nécessaire de manipuler de grands ensembles d'informations.

I.2 Problématique : comment maîtriser un espace informationnel ?

Maîtriser ses informations devient alors le challenge incontournable auquel doit faire face toute organisation indépendamment de sa taille.

« Crises, accélération des mutations économiques et sociales, globalisation. Le monde est de plus en plus complexe. Le dirigeant d'entreprise rencontre des difficultés croissantes pour garder le cap. Ses traditionnels instruments de navigation deviennent obsolètes. Entre tempête et brouillard, la maîtrise de l'information devient un enjeu stratégique pour éviter les écueils. » Jean-Eudes Tesson et Pierre Fonlupt dans [Corman & Ingargiola, 2005]

Nous nommons « *espace informationnel* » l'ensemble des informations manipulées par une organisation. Pour chaque organisation, cet espace est très vaste. Il va des informations collectées (comme les informations sur les collaborateurs, les clients, les fournisseurs, les partenaires, les concurrents, les marchés, etc.) aux savoir-faire (comme les descriptions des outils et des procédés, la documentation scientifique et technique, les brevets, etc.).

L'avenir d'une organisation passe donc par la maîtrise de son espace informationnel. Cet objectif n'est possible que si elle dispose des moyens nécessaires pour en avoir une vision organisée, globale et précise. En effet, avec les travaux portant sur la gestion des connaissances, le problème n'est plus l'archivage ou l'indexation des informations, mais celui de leur accès à travers une appréhension globale reposant sur les connaissances de l'organisation.

L'objectif final est de permettre aux collaborateurs d'appréhender l'espace informationnel de leur organisation d'une manière globale et de se focaliser sur ce qui est important pour leur activité.

¹ Nous considérons les connaissances d'une organisation comme l'ensemble de son patrimoine intellectuel caractérisé par ses connaissances explicites (les documents, projets, retours d'expériences, etc...) et les connaissances tacites détenues par les collaborateurs (les expériences, le savoir-faire, etc. ...).

I.3 Notre approche : cartographe pour maîtriser

La maîtrise d'un espace est une problématique universelle qui préoccupe l'homme depuis son origine. La cartographie résulte de cette problématique. Cartographier un espace permet de le visualiser en une carte pour mieux l'appréhender¹.

« Une image vaut mille mots ». Cette constatation sur la valeur de la visualisation est d'autant plus vraie si l'on travaille avec un grand nombre d'informations. La synthèse des travaux sur la cartographie (qu'elle soit géographique ou thématique) et sur la visualisation permet de considérer les cartes comme un médium idéal entre un grand nombre d'informations et l'esprit [Ware, 2000]. Chaque carte joue alors le rôle de support externe de la pensée pour amplifier la cognition² de ses utilisateurs.

Face à ce constat, notre approche consiste à étudier les différents paradigmes de la cartographie et de les appliquer aux informations manipulées par les organisations pour construire des cartes. Nous pensons que la carte peut fournir un support efficace pour l'aide à la décision (sous réserve de fournir des cartes adaptées) et ainsi aider les collaborateurs à appréhender leur espace informationnel pour mieux le maîtriser.

Nous pouvons donc résumer notre approche comme étant la cartographie de l'espace informationnel des organisations pour mieux le maîtriser.

Nous nommons « cartographie de données abstraites », le recours aux paradigmes de cartographie pour représenter un espace informationnel. Bien que cette nomination soit discutable, nous la choisissons en référence à la « visualisation de données abstraites » définie par la communauté de la visualisation d'informations.

I.4 Point de départ : besoins des utilisateurs

Le nombre de paradigmes de cartographie est très important ; par conséquent il est possible de construire un nombre quasiment infini de cartes différentes d'un même espace de données. Dans un tel contexte, comment construire une carte qui aidera effectivement ses utilisateurs à appréhender leur espace informationnel ?

« Devant la masse toujours croissante des données, la cartographie offre des possibilités de communiquer synthétiquement et dégager les informations stratégiques utiles aux décideurs. Seulement, comment passer des données brutes à la carte ? » Pierre Mongin (préface de [Poidevin, 1999]).

De plus, une mauvaise carte peut véhiculer une information ambiguë ou erronée et son utilisateur peut se retrouver pénalisé dans ses prises de décisions. Pour éviter ce genre de

¹ D'après le Trésor de la Langue Française, l'appréhension est la faculté de saisir par l'intelligence et plus particulièrement, toute opération intellectuelle relativement simple ou immédiate, soit de jugement, soit de mémoire, soit d'imagination, considérée comme s'appliquant à un contenu distinct de l'opération elle-même.

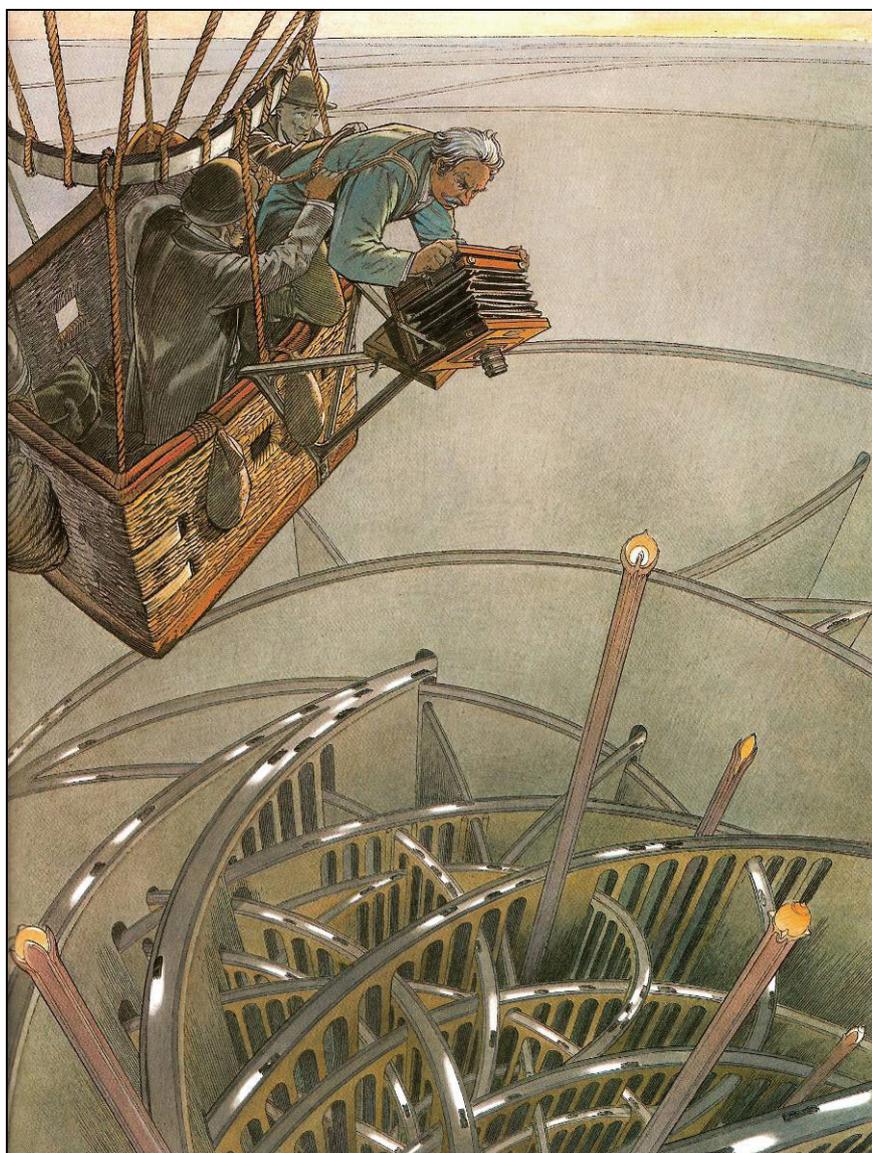
² Fonction complexe multiple regroupant l'ensemble des activités mentales (pensée, perception, action, volonté, mémorisation, rappel, apprentissage) impliquées dans la relation de l'être humain avec son environnement et qui lui permettent d'acquérir et de manipuler des connaissances (associations, rétroaction, traitement de l'information, résolution de problèmes, prise de décision etc.). (Le Grand Dictionnaire Terminologique)

mésaventure, nous avons identifié plusieurs besoins que la cartographie devait satisfaire pour répondre à notre problématique d'appréhension d'un espace informationnel :

- > **Naviguer selon la sémantique du domaine** entendue comme l'ensemble des concepts permettant d'appréhender le domaine : pour leurs activités, les collaborateurs doivent naviguer dans l'espace informationnel de leur organisation (comme lors d'une recherche d'informations). Pour faciliter cette navigation, il est nécessaire que les cartes véhiculent la sémantique du domaine. Si la cartographie respecte cette sémantique, les utilisateurs peuvent alors la comprendre, l'assimiler et l'exploiter. Utiliser une carte qui exploite la sémantique du domaine pour naviguer permet aux utilisateurs d'appréhender plus facilement l'espace informationnel.
- > **Proposer une vision à plusieurs échelles** : permettre aux utilisateurs de maîtriser un ensemble d'informations nécessite de leur fournir les moyens pour l'appréhender dans sa globalité mais aussi dans ses particularités. C'est pourquoi, une carte doit offrir simultanément une vision globale et synthétique de l'espace informationnel, et une vision particulière et détaillée de ce même espace.
- > **Proposer une carte adaptée à l'utilisateur** : tous les collaborateurs n'ont pas la même activité et le même niveau d'expertise. Chaque carte doit donc être adaptée à ses utilisateurs. De plus, une carte étant une représentation de l'espace informationnel, elle véhicule une signification. Il est alors nécessaire de bien choisir et adapter cette représentation pour maîtriser la signification qui sera finalement perçue.

L'ensemble de ces besoins est lié à la nécessité pour l'utilisateur de naviguer dans un espace informationnel souvent important. Nous nommons « **cartographie sémantique** » la « cartographie de données abstraites » de l'espace informationnel d'une organisation répondant à ces besoins. Ainsi, elle devient une activité essentielle à la gestion des connaissances, permettant de tirer parti de toute la richesse des informations de l'organisation.

Partie 2 : **État de l'art**



*« Qu'une bonne carte est chose merveilleuse, on y regarde le monde
comme vu d'un autre monde grâce à l'art du dessin. »*

Samuel Van Hoogstraten

Rappel du plan

2.1 La cartographie	25
2.1.1 Un monde de cartes	25
2.1.2 Quelques définitions	29
2.1.3 Principe de l'information cartographique.....	30
2.2 La cartographie de données abstraites	41
2.2.1 Une activité pluridisciplinaire	41
2.2.2 Les approches courantes.....	42
2.2.3 Un processus en niveaux	52
2.2.4 Les principaux paradigmes.....	54
2.3 Synthèse & critiques	90

Cette partie décrit un état de l'art du domaine de la cartographie de données abstraites en débutant par la présentation succincte de la cartographie en générale. Ensuite, elle introduit la cartographie de données abstraites avec ses deux communautés et son ensemble de paradigmes pour cartographier un espace informationnel selon nos besoins. Enfin, cette partie se termine par une synthèse critique de l'existant, permettant d'introduire nos différentes propositions.

2.1 La cartographie

Pour aider à comprendre la fonction et les enjeux de la cartographie sémantique, cette section retrace l'évolution de la cartographie depuis ses origines à nos jours. Il introduit la cartographie de données abstraites comme une nouvelle branche dans l'arbre de l'évolution de la cartographie.

2.1.1 Un monde de cartes

Le monde des connaissances n'est pas très différent de notre monde physique. Il est aux organisations ce que le monde physique est aux hommes : vaste, dynamique et omniprésent.

Que le contenu soit géographique ou abstrait, l'homme doit sans cesse évoluer dans des espaces différents, prendre des décisions et interagir avec eux. Pour y parvenir, il doit les comprendre et les maîtriser. Mais comment appréhender et comprendre des espaces aussi complexes ? Comment en avoir une vision adaptée aux besoins, aux objectifs et à l'activité de chacun ? Comment en avoir une vision qui soit tout à la fois globale et locale, générale et particulière, schématisée et précise, universelle et dédiée ?

Cette problématique est à l'origine de la cartographie : le passage par une représentation graphique adaptée d'un territoire. La citation suivante, pourtant extraite d'un livre sur la cartographie géographique [Poidevin, 1999], rejoint notre problématique et l'enjeu de la cartographie en générale :

« [...] une problématique à laquelle sont confrontés quotidiennement beaucoup de professionnels : celle de comprendre et de gérer leurs territoires grâce à la cartographie. L'enjeu est grand : meilleurs seront les documents cartographiques, plus sûres seront les décisions qui en émanent. »

Elle s'applique parfaitement à notre contexte de la cartographie de données abstraites.

a. La cartographie antique

Pour la plupart des civilisations, les premiers signes de l'utilisation de cartes coïncident avec l'invention de l'écriture [Rouleau, 2000]. Au début, ces cartes sont des esquisses gravées qui ont fonction d'aide-mémoire. Il ne s'agit alors que de représentations d'éléments constitutifs du territoire, de l'environnement « proche », comme les hommes le percevaient.

La plus ancienne carte que l'on ait retrouvé jusqu'à présent, est cette tablette babylonienne appelée "Carte babylonienne du monde" :



Figure 3 - Carte babylonienne du monde, composée probablement entre 700 et 500 ans avant Jésus-Christ (British Museum, Londres, Salle 55, Later Mesopotamia, case 15, no. 27).

On y voit un cercle, l'Océan, autour duquel se trouvent des régions fabuleuses décrites dans le texte cunéiforme. Au centre se trouve le monde connu, orienté avec l'Ouest en haut. La case supérieure porte le nom de Babylone. Les lignes verticales qui traversent Babylone représentent l'Euphrate. Cette carte fut composée probablement entre 700 et 500 ans avant Jésus-Christ.

Rapidement, les cartes sont devenues l'outil des navigateurs qui y reportent sur des pierres, du bronze ou du papyrus les éléments utiles de la configuration des côtes, l'emplacement des ports, les récifs à contourner, etc. Cette période de l'évolution de la cartographie est marquée par de grands scientifiques grecs comme Ératosthène (275-194 avant Jésus-Christ), Hipparque (190-125 avant Jésus-Christ) et Ptolémée (90-168) [Poidevin, 1999]. Ils avaient en commun la volonté de cartographier la terre et par là même, de démontrer sa sphéricité.

Au fil du temps et des besoins, la cartographie s'efforce de plus en plus de reproduire le réel pour fournir des aides dans la vie quotidienne. L'exemple des plans cadastraux des Égyptiens au temps de Ramsès II est révélateur de cette évolution.

b. La cartographie topographique

Le moyen-âge correspond à l'apparition d'une cartographie de qualité (plus soignée, plus précise, plus détaillée) grâce aux cartographes arabes. Puis, à partir du XIII^e siècle et pendant deux cent cinquante ans, se développe une cartographie « utile ». Il en résulte la création de cartes très spécifiques comme les cartes nautiques. La cartographie s'oriente de façon décisive vers la « couverture » topographique descriptive, la plus complète et la plus exacte possible de la surface du globe : la représentation du visible.

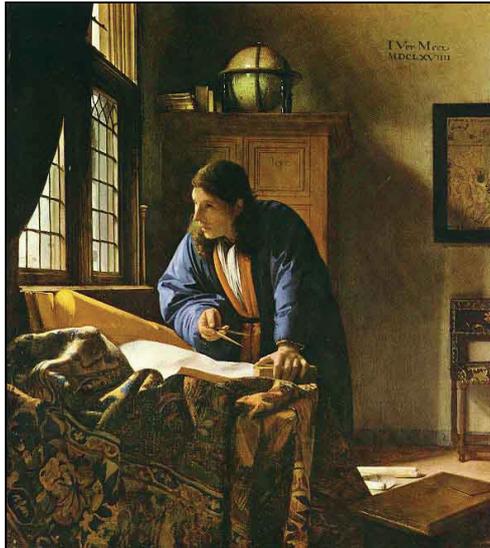


Figure 4 - « Le géographe » de Johannes Vermeer.

Les XVIIe, XVIIIe et XIXe siècles marquent la naissance de la cartographie topographique dont l'un des maillons essentiels est la « Carte de Cassini » (voir figure ci-dessous) [Poidevin, 1999]. Cette carte a pour objectif de représenter le Royaume de France et les travaux de levée furent à l'initiative de Louis XV.

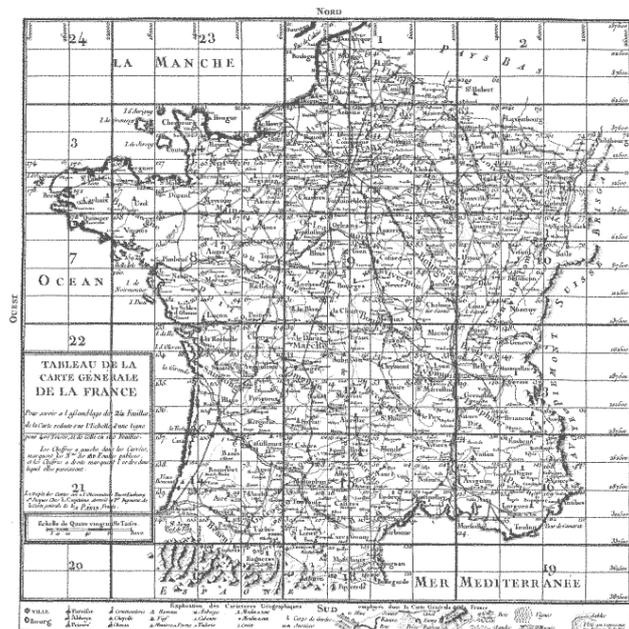


Figure 5 - Carte de Cassini (source IGN).

Elle sera réalisée en 55 ans (1760 - 1815) par quatre générations de la famille Cassini et elle servira de référence aux cartographies des principales nations européennes pendant la première moitié du XIXe siècle.

c. La cartographie thématique

Dès le XIXe siècle apparaissent de nouvelles cartes avec des données à visualiser qui ne sont plus des localisations géographiques mais des données qualitatives et quantitatives. Par

exemple, ces données peuvent concerner les composantes de l'occupation agricole des sols, les premiers recensements nationaux ou bien encore des données statistiques. Il s'agit alors de la naissance de la cartographie thématique.

La cartographie thématique a pour objectif de donner une représentation conventionnelle de tous phénomènes à distribution spatiale et de leurs corrélations. « La rédaction des cartes thématiques au niveau du levé et jusqu'à la conception de la maquette est avant tout le fait des divers spécialistes des thèmes cartographiés. Il s'agit en effet de traiter une documentation puisée aux sources mêmes des sciences, des disciplines ou des techniques servies. »¹

La « Carte figurative des pertes successives en hommes de l'Armée française dans la campagne de Russie de 1812-1813 » est un des meilleurs exemples de carte thématique. Elle a été décrite comme étant « le meilleur graphique jamais produit » [Tufte, 1986], un de ceux qui semblent « défier la plume de l'historien par sa brutale éloquence » [Marey, 1878]. Elle représente le parcours des troupes de Napoléon (données géographiques) ainsi que leur importance (données quantitatives).

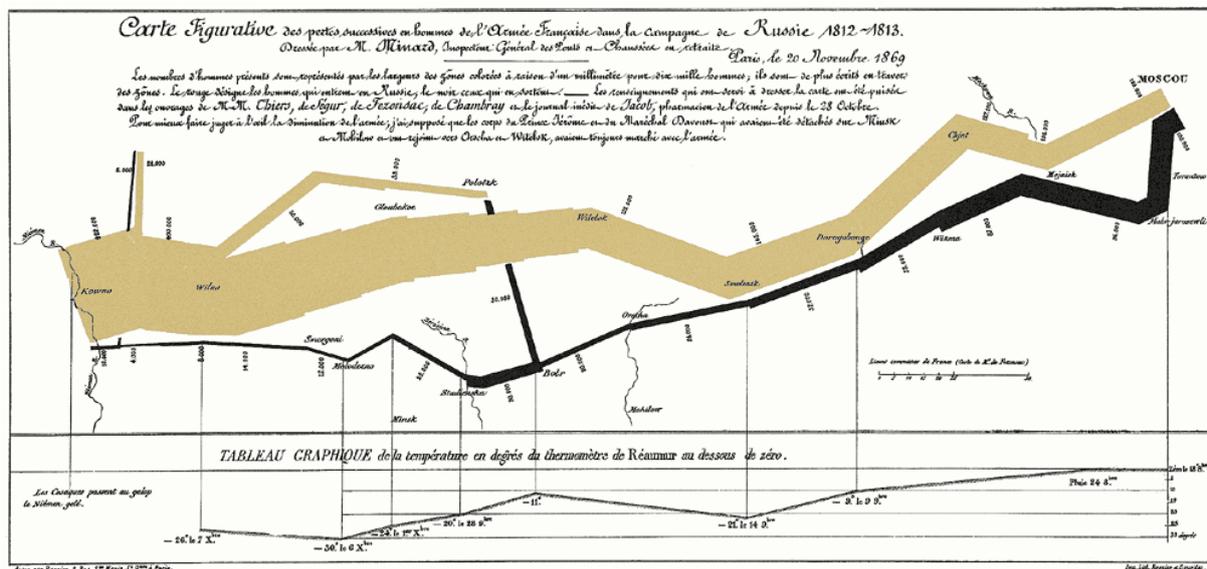


Figure 6 - « Carte figurative des pertes successives en hommes de l'Armée française dans la campagne de Russie de 1812-1813 » de Charles Joseph Minard (1869).

Le développement rapide de la cartographie thématique est une évolution importante dans l'histoire de la cartographie. Il traduit l'intérêt croissant du public pour l'usage de l'expression cartographique et témoigne d'une maîtrise de plus en plus sûre des techniques dont disposent les cartographes. Il implique cependant une telle diversité de recueil des données et de présentation des résultats qu'il échappe pratiquement à tout effort de normalisation.

La cartographie voit donc apparaître une nouvelle expression visuelle, une nouvelle imagerie cartographique qui ne relève plus uniquement de données géographiques (des localisations, des distances, des altitudes) mais aussi de données de plus en plus abstraites : « le passage de nouveaux concepts à une nouvelle figuration » [Rouleau, 2000].

¹ D'après "Universalis Multimédia" 11ème édition : Encyclopædia Universalis, 2005

Grâce à la prise de conscience de ses qualités pour l'aide à la décision, la cartographie connaît un développement prodigieux. Son évolution débute par le traitement de données physiques (essentiellement géographiques) pour tendre vers le traitement de données abstraites. Elle devient un enjeu primordial dans de nombreux domaines ayant un lien avec la gestion d'espaces (données géographiques ou abstraites).

2.1.2 Quelques définitions

Nous savons que la cartographie permet de produire des cartes mais, sur quels principes repose la cartographie et que contient une carte ? Pour commencer à répondre à cette question, nous avons choisi de débiter par l'étude des définitions associées à la cartographie.

Étant donné que la cartographie est très répandue depuis de nombreux siècles, il en existe de nombreuses définitions. Comme référence, nous avons sélectionné celle proposée par le Comité Français de Cartographie¹ :

« La **cartographie** est l'ensemble des études et des opérations scientifiques, artistiques et techniques, intervenant à partir des résultats d'opérations directes ou de l'exploitation d'une documentation, en vue de l'élaboration et de l'établissement de cartes, plans et autres modes d'expression, ainsi que dans leur utilisation. »
Le comité Français de Cartographie

Tout comme la cartographie, nous avons choisi la définition suivante pour la carte :

« La **carte** est une représentation géométrique conventionnelle, généralement plane, en positions relatives, de phénomènes concrets ou abstraits, localisables dans l'espace ; c'est aussi un document portant cette représentation ou une partie de cette représentation sous forme d'une figure manuscrite, imprimée ou réalisée par tout autre moyen.»
Le comité Français de Cartographie

Nous avons retenu ces définitions pour plusieurs raisons. Tout d'abord, bien qu'elles soient rédigées par une communauté de géographes, elles ne contraignent pas la nature des données cartographiées. Elles ne limitent donc pas le champ de la cartographie à des données géographiques. Ensuite, nous les avons choisies car elles sont consensuelles (au moins pour la communauté des géographes) et car elles sont les résultats d'une longue évolution de définitions depuis l'origine de la cartographie. Elles en reprennent les justesses et en évitent les imprécisions.

Le **cartographe** conçoit et dresse les cartes, il cartographie :

« Le **cartographe** [est] un opérateur capable de maîtriser les concepts et les taxons de la discipline qu'il se charge d'illustrer ».
Encyclopædia Universalis².

Ces quelques définitions nous éclairent un peu plus sur la cartographie. Néanmoins, elles ne nous expliquent pas le principe de l'information cartographique. C'est le rôle de la section suivante.

¹ <http://www.lecfc.freesurf.fr>

² D'après "Universalis Multimédia," | 1^{ème} édition : Encyclopædia Universalis, 2005

2.1.3 Principe de l'information cartographique

Dans ses travaux sur la cartographie géographique, Bernard Rouleau pose trois questions fondamentales pour comprendre la cartographie et plus particulièrement les cartes [Rouleau, 2000] :

« En quoi consiste une carte ? Quel est le contenu d'une carte ? À quoi sert une carte ? ».

Les paragraphes suivants reprennent chacune de ces questions en leur associant la réponse donnée par l'auteur (expert en cartographie géographique). Étant donné que l'on s'inspire de la cartographie géographique pour l'appliquer à des données abstraites, nous avons généralisé les réponses à la cartographie de données non uniquement géographiques :

En quoi consiste une carte ?

Il s'agit toujours d'un **document graphique** donc visuel. À ce titre, toutes les cartes suivent les règles et les contraintes de la perception visuelle [Tufté, 1986; Ware, 2000]. Par conséquent elles obéissent à la sémiotique (théorie générale des signes).

Quel est le contenu d'une carte ?

Une carte représente toujours un espace d'informations. Ces informations peuvent être abstraites (comme les données d'une organisation) ou scientifiques (c'est-à-dire issues de mesures physiques comme des données géographiques) [Card et al., 1999b]. Les informations sont *visualisées* par des **représentations graphiques** (les composants pour les géographes) et *sélectionnées* (une carte ne fait apparaître qu'une catégorie d'informations : son *thème* [Rouleau, 2000]). Ainsi, une carte permet de connaître la nature, les relations, la localisation ou l'importance des données ou des phénomènes qui composent l'espace cartographié. Par conséquent, les représentations graphiques des données et leurs relations forment l'image cartographique : le *contenu* de la carte [Rouleau, 2000].

À quoi sert une carte ?

Une carte n'est pas une simple image artistique d'un espace informationnel. L'objectif avant tout d'une carte est de **transmettre une information** : « Le message de la carte » [Rouleau, 2000]. Cependant, il arrive que la carte soit utilisée comme méthode de recherche lorsqu'elle apporte elle-même une information nouvelle non encore prévue ou pressentie. La carte devient alors un instrument de **transmission de messages mais aussi un instrument de résolution**.

Au fil des siècles, la cartographie est devenue un art d'expression mais aussi, un outil d'analyse, d'aide à la décision et de communication. Elle a évolué de données topographiques et sommaires sur un lieu donné, à des données abstraites appartenant à une organisation ou un domaine, en passant par la couverture complète du globe terrestre et du ciel. La cartographie parfait ses techniques pour devenir un outil généralisé d'aide à la décision.

Pour comprendre le principe intrinsèque de la cartographie, il est important de comprendre le rôle joué par le « monde externe » dans le raisonnement.

a. La cognition externe

Des travaux sur le raisonnement humain ont permis de mettre en évidence que lorsque l'homme a besoin de raisonner, à un certain niveau de difficultés et pour une certaine quantité d'informations, il utilise invariablement des supports externes pour l'assister comme un tableau, une feuille de papier ou bien encore une carte. Ce principe est nommée la cognition externe (« external cognition ») [Scaife & Rogers, 1996].

La citation suivante est extraite du livre « Things that make us smart : defending human attributes in the age of the machine » de Donald Norman [Norman, 1993] ; elle exprime sa vision de la cognition externe qu'il nomme « aides externes » :

“The power of the unaided mind is highly overrated. Without external aids, memory, thought, and reasoning are all constrained. But human intelligence is highly flexible and adaptive, superb at inventing procedures and objects that overcome its own limits. The real powers come from devising external aids that enhance cognitive abilities. How have we increased memory, thought, and reasoning? By the inventions of external aids: It is things that make us smart.” Donald Norman [Norman, 1993]

Pour illustrer ce phénomène, prenons un exemple simple mais aussi très révélateur de l'efficacité et du rôle de la cognition externe : la multiplication. Pour un calcul simple, la multiplication peut « se faire de tête » ; on parle alors de cognition interne. Par exemple, si on veut multiplier 2 par 2, on peut effectuer le calcul mentalement.

Mais lorsque les valeurs multipliées sont plus grandes, il devient alors très difficile d'effectuer ce calcul « de tête » uniquement. Par exemple, si on veut multiplier 28 par 43, le calcul mental est plus difficile.

Pour y parvenir simplement, la solution la plus commune est d'utiliser un papier et un crayon et de « poser » l'opération :

A handwritten multiplication problem on a grid background. The numbers are written in blue ink. The problem is 28 multiplied by 43. The calculation is shown in a standard 'positional' format with horizontal lines separating the partial products. The first partial product is 184 (28 times 4), and the second is 1120 (28 times 40). The final result is 1204.

$$\begin{array}{r} 28 \\ \times 43 \\ \hline 184 \\ 1120 \\ \hline 1204 \end{array}$$

Figure 7 - Poser une opération : un exemple de cognition externe.

Cette représentation est un exemple de cognition externe dont le gain d'efficacité est évident.

La question que l'on se pose alors est de savoir pourquoi une aide externe (comme le papier ou une carte) facilite-t-elle le calcul et le raisonnement ?

Si on reprend notre dernier exemple de calcul mental, la difficulté réside dans la mémorisation des retenues et des différents résultats intermédiaires. Intuitivement, pour contourner cette difficulté liée à ce manque de « mémoire », nous utilisons une feuille de papier comme support externe et on y représente graphiquement l'opération. Le papier joue alors le rôle d'extension de la mémoire [Card et al., 1999b].

Le travail intellectuel (ou tâche cognitive) exploite un ensemble de connaissances qui sont représentées dans l'esprit de l'homme. Cette représentation est qualifiée de « représentation interne » [Zhang & Norman, 1994]. L'utilisation d'un support externe (comme des symboles ou des objets) permet de « compléter » cette représentation interne et ainsi décharger un peu la mémoire. Le support externe est qualifié de « représentation externe ».

La représentation externe est perçue, analysée puis traitée par le système perceptif qui en extrait les informations. Tout ce traitement correspond au mécanisme de la vision.

Dans le cadre d'une activité utilisant un support externe, le traitement d'une tâche est un processus distribué entre des représentations internes et des représentations externes. Ce mécanisme est nommé « raisonnement distribué » [Zhang & Norman, 1994]. Il est composé de deux types de phases permettant au cerveau de passer d'une représentation interne à une représentation externe (externalisation) et inversement (mémorisation).

Voici un schéma de synthèse des deux phases du raisonnement distribué :

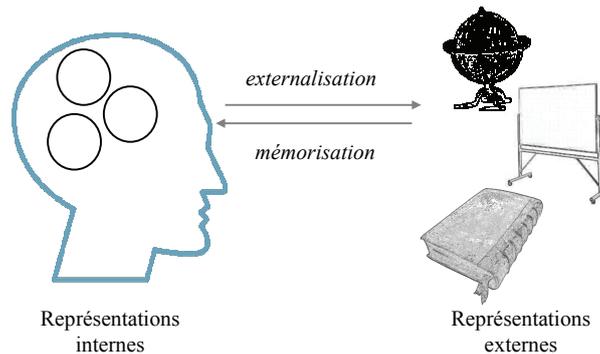


Figure 8 - La cognition externe : phases du raisonnement distribué.

Une représentation externe peut être transformée en représentation interne par « mémorisation » [Zhang, 1997]. La mémorisation n'est pas obligatoire si la représentation externe est toujours accessible et elle est impossible si les représentations externes sont trop complexes. Une représentation interne peut aussi être transformée en représentation externe par « externalisation ». L'externalisation peut être bénéfique si l'utilisation d'une représentation externe peut compenser le coût du processus d'externalisation.

Les représentations externes jouent donc le rôle d'aide-mémoire en permettant d'une part, d'étendre la mémoire de travail et d'accéder à des informations non mémorisées et d'autre part, en permettant de partager des connaissances. De plus, l'utilisation de représentations externes permet aussi d'accéder à des connaissances et compétences inaccessibles par des représentations internes [Zhang, 1997].

Nous savons donc que les supports externes permettent d'assister l'homme dans ses tâches cognitives. De plus, nous avons vu que les supports externes pouvaient être de natures très diverses. Le problème est donc de savoir **quelle représentation externe choisir ?**

Les représentations externes facilitent le travail cognitif seulement si la représentation est bien choisie. En effet, deux représentations d'une même structure formelle peuvent engendrer des comportements cognitifs très différents [Zhang & Norman, 1994].

Une illustration évidente est la représentation des nombres. Tout le monde s'accorde à dire que pour calculer, les chiffres arabes sont plus efficaces que les chiffres romains (par exemple, 73×27 est plus facile que $LXXIII \times XXVII$). Pourtant, ces deux types de numérotations expriment les mêmes entités, des nombres. Des représentations différentes d'une même structure formelle peuvent donc avoir des impacts très différents sur la difficulté d'une tâche.

b. La visualisation

Nous avons vu que face à un grand nombre d'informations, l'homme utilise naturellement des représentations externes pour l'aider. Il nous reste alors à déterminer le type de support externe à utiliser.

Pour répondre à cette question, nous pouvons tout d'abord remarquer que face à un grand nombre d'informations, l'homme privilégie généralement les représentations visuelles graphiques comme support externe sans être conscient de la cause qui le pousse à y recourir.

C'est le cas avec l'utilisation de cartes géographiques pour appréhender un territoire ou d'un graphique pour représenter des ensembles de valeurs numériques.

Pourquoi utiliser des représentations visuelles graphiques ?

Ce choix est justifié par les capacités visuelles naturelles de l'homme. Nous sommes capables d'assimiler quasi instantanément et sans effort un grand nombre d'informations représentées graphiquement. Par exemple, le recours à une carte routière pour déterminer un itinéraire illustre cette pratique. L'œil perçoit entre autres, immédiatement la position relative des villes sur la carte routière.

Les visualisations profitent des caractéristiques du système de traitement cognitif humain. D'après Colin Ware, les visualisations offrent de nombreux avantages dont les principaux sont les suivants [Ware, 2000; Ware, 2005] :

- > Les visualisations graphiques ont le pouvoir de permettre à l'homme de manipuler des structures bien plus complexes représentées par une visualisation (représentations externes) que dans la mémoire de travail visuelle et verbale (représentations internes).
- > Elles permettent aussi de percevoir l'émergence de propriétés dans les données cartographiées qui ne sont pas anticipées.
- > Les visualisations permettent de mettre en évidence des problèmes dans les données, dans leur collecte. Avec une visualisation appropriée, les erreurs dans les données sont rapidement perceptibles.
- > La visualisation permet de percevoir simultanément des propriétés à grande et à petite échelle sur les données.
- > La visualisation facilite la formation d'hypothèses sur les données.

Durant un processus d'acquisition ou de résolution de problèmes (raisonnement), une visualisation peut aider l'utilisateur à dépasser les problèmes qui sont dus aux limitations de la mémoire de travail : capacité de mémorisation et durée pour mémoriser [Keller & Tergan, 2005]. Comparé à une description textuelle équivalente, un diagramme peut permettre aux utilisateurs d'éviter de traiter explicitement l'information car ils peuvent l'extraire « d'un coup d'œil » [Keller & Tergan, 2005]. Les visualisations peuvent augmenter nos capacités de traitement en visualisant des relations abstraites entre les éléments visualisés [Scaife & Rogers, 1996].

Les visualisations peuvent être considérées comme des outils capables d'assister l'homme dans son raisonnement. Elles permettent d'amplifier la cognition en jouant le rôle de support de la pensée (raisonnement distribué).

Jacques Bertin, un géographe contemporain de grande renommée, nomme le recours à des visualisations graphiques pour raisonner le « traitement graphique de l'information » ou « la graphique » [Bertin, 1977].

c. La graphique

Dès 1950, Jacques Bertin a travaillé sur les aspects théoriques de la construction d'une carte et plus particulièrement, sur les principes de l'information cartographique. C'est ainsi que quelques années plus tard, il a défini les notions de « graphique » et de « sémiologie graphique » [Bertin, 1977; Bertin, 1999 (1^{re} éd. 1969)].

i. La sémiologie graphique

En tant que moyen d'expression, la carte utilise un langage pour coder l'information du message à véhiculer. Ce langage a la particularité d'être graphique et de mettre en œuvre des *structures visuelles* [André, 1980]. Ces structures sont perçues par le sens de la vue. Les mécanismes de la perception visuelle sont présentés synthétiquement en annexes (voir en page 237).

Le message véhiculé par la carte est donc codé par un ensemble de structures visuelles. Chaque structure visuelle peut appartenir à la famille des *dessins* ou à la famille des *écritures*. La différence entre un dessin et une écriture réside dans le type d'accès à l'information codée par la structure. Dans le cas du dessin, l'information représentée est accessible directement. Alors que dans le cas de l'écriture, l'information est dénotée par un mot qui est représenté (écrit). Pour percevoir l'information, il faut tout d'abord percevoir la forme écrite qui donne accès aux mots et enfin, accéder à l'information. Cette particularité explique pourquoi la carte est souvent qualifiée de document « immédiat ».

La famille des dessins se ramifie encore en deux sous-familles : le « *dessin d'art* » (ou « *graphisme* ») et « *la graphique* » [Bertin, 1977]. Une représentation graphique (un graphique ou un graphisme) est la transcription de concepts ou d'idées sous la forme codée d'un système de structures visuelles (les signes). Plus particulièrement, la graphique correspond à la construction d'images à partir d'une grammaire qui s'appuie sur les lois de la perception visuelle, perception universelle. L'étude et la mise en œuvre de ces règles –nommées « sémiologie graphique » par Jacques Bertin [Bertin, 1999 (1^{re} éd. 1969)]– ont permis de comprendre et d'exposer la différence fondamentale entre graphisme et graphique : leur domaine de significations.

D'après Jacques Bertin, la perception d'un graphisme permet plusieurs significations : le dessin d'art figuratif (comme une photographie aérienne) permet une signification par récepteur (personne qui perçoit le dessin) ; alors que le dessin d'art non figuratif (comme un tableau abstrait) est dit pansémique¹. Il existe donc une infinité d'interprétations et de significations possibles pour un même récepteur. Le processus de la perception de l'image se traduit par la question : « Que signifie l'ensemble de ces éléments ? ». Pour Jacques Bertin, le travail de lecture d'un graphisme se situe alors « entre le signe et sa signification ».

Le graphique (produit de la graphique) est quant à lui un ensemble de structures visuelles dont la signification a été définie au préalable. Le processus de perception d'un graphique se traduit alors par la question : « Étant donné la signification de ces éléments, quelles sont les relations qui s'établissent entre eux ? » Pour Jacques Bertin, le graphique est monosémique et le travail de lecture se situe alors « entre les significations ».

¹ « Pansémique » qualifie un message ouvert à toutes les interprétations. Le Grand Dictionnaire Terminologique

Cette étude des significations des dessins donne l'avantage à la graphique par rapport au graphisme : utiliser un système monosémique permet de réduire la confusion. Les récepteurs d'un graphique partagent la même signification pour les structures visuelles. Ce consensus permet de « discuter de l'assemblage des signes [structures visuelles] et d'enchaîner les propositions dans une succession d'évidences, succession qui peut alors devenir indiscutable c'est-à-dire logique » [Bertin, 1999 (1re éd. 1969)]. Pour Jacques Bertin, graphique et mathématique sont semblables sur ce point.

Le schéma suivant synthétise les différents systèmes d'expression graphique :

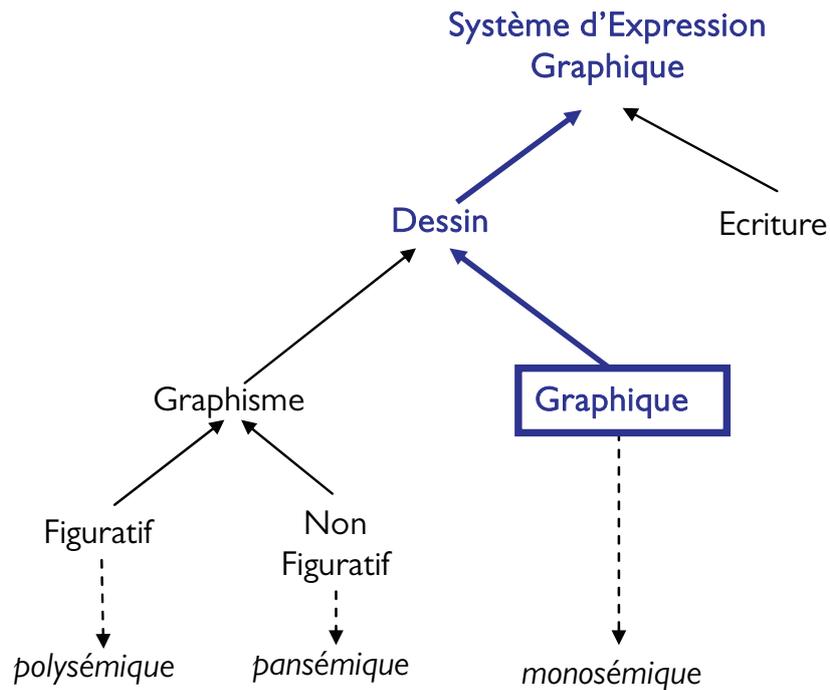


Figure 9 - La graphique par rapport aux autres systèmes d'expression graphique.

Voici un tableau récapitulatif des domaines de signification liés aux dessins :

Type de dessin		Signification des signes	Description	Exemple
Graphisme	Image non-figurative	Pansémique	Le système s'ouvre à toute signification.	Tableau d'art
	Image figurative	Polysémique	Le système a pour objectif de définir un concept ou une idée mais les interprétations peuvent diverger.	Photographie aérienne.
Graphique		Monosémique	Transcriptions de relations entre des concepts préalablement définis.	Un organigramme

Tableau 1 - Domaines de signification des dessins.

L'étude et l'analyse des significations des différents systèmes d'expression graphique ne doivent pas faire oublier l'importance du caractère esthétique des cartes [Tractinsky, 1997]. En effet, le but d'un graphique est avant tout d'être *efficace*¹ et pour y parvenir, il doit être correctement construit, mais il doit aussi être esthétique. D'après Albert André, une image esthétique, sans grande valeur symbolique retient l'attention d'un lecteur, tandis qu'une image laide manque son effet [André, 1980]. C'est l'esthétique qui, au premier abord, commande l'efficacité d'une image. Une image esthétique est celle qui est agréable à regarder et que la mémoire enregistre.

ii. Les structures visuelles et les variables visuelles

« La graphique est un système de signes [structures visuelles] qui permet de transcrire les relations de différence, d'ordre ou de proportionnalité existant entre des données qualitatives ou quantitatives » [Bonin, 1997].

Comme nous l'avons décrit dans la section précédente, une image peut être décomposée en structures visuelles. Chaque structure visuelle appartient à une des quatre familles suivantes (Jacques Bertin parle alors d'« implantation » des signes) [Bertin, 1977; Bertin, 1999 (1re éd. 1969); Card et al., 1999b] :

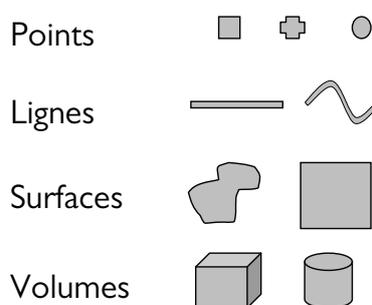


Figure 10 - Familles de structures visuelles.

L'appartenance d'une structure visuelle à une famille est fonction de sa dimensionnalité.

Pour coder des informations, la graphique fait varier certaines propriétés graphiques de ces structures visuelles (par exemple la forme). Les variations possibles sur les structures visuelles sont regroupées par type et sont nommées « variables visuelles » (ou « variables rétinienne ») par Jacques Bertin [Bertin, 1977; Bertin, 1999 (1re éd. 1969)].

Le schéma ci-dessous reprend les variables visuelles définies par Jacques Bertin :

¹ « qui produit l'effet qu'on attend », dictionnaire Robert.

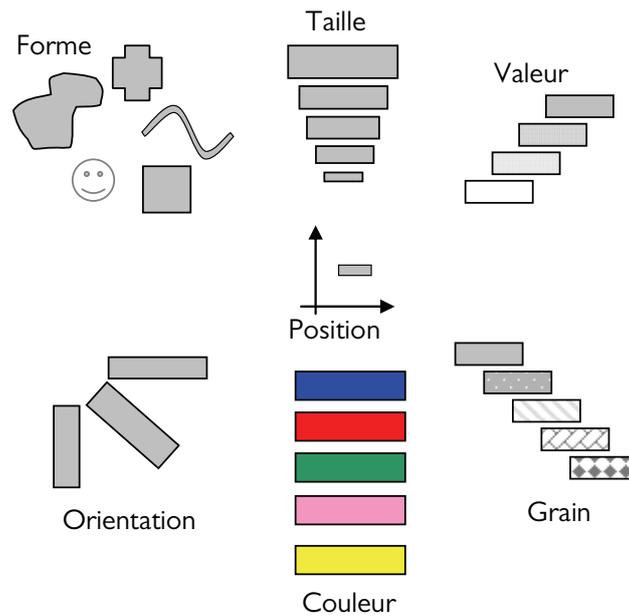


Figure 11 - Variables visuelles définies par Jacques Bertin.

Au fil des évolutions des moyens de cartographie, le nombre de ces variables a évolué avec l'identification de nouvelles variables¹. Elles sont généralement regroupées en deux familles : les variables statiques et les variables dynamiques. Voici un tableau de synthèse des principales variables visuelles :

Variables Statiques	Forme Taille Valeur Grain Couleur Orientation Saturation de couleurs Aspect de la texture Pattern Finesse des détails Luminance de l'écran Ombre
Variables Dynamiques	Vitesse de mouvement Direction de mouvement Fréquence de clignotement Phase de clignotement Disparité binoculaire

Tableau 2 - Synthèse des principales variables visuelles.

iii. Interactions entre les structures visuelles : la théorie de la Gestalt

Les philosophes grecs pensaient que la nature des choses était absolue et qu'elle n'était pas liée au contexte de celles-ci. Mais vers 1890, la théorie de la Gestalt est apparue comme une

¹ Un important travail de synthèse sur ce sujet, nommé "Extension des variables graphiques définies par Bertin", est disponible sur le site de l'Institut National de l'Audiovisuel (INA) à l'adresse suivante : <http://www.ina.fr/recherche/theme/vie/interf/extension.fr.html>.

réaction à ce courant de pensée. Des adeptes de cette nouvelle vision étaient intrigués par la manière dont notre cerveau percevait des ensembles à partir d'éléments inachevés. Plus particulièrement, ils ont mis en évidence que le contexte est très important dans la perception visuelle [Guillaume].

Une carte contient un ensemble de structures visuelles qui codent les éléments de l'espace informationnel (les informations à cartographier). Par conséquent et d'après la théorie de la Gestalt, les structures visuelles ne sont pas indépendantes les unes des autres et le contexte est important pour la signification qui leur est attribuée. La perception visuelle s'attache à reconnaître des modèles dans un ensemble de structures visuelles.

De cette théorie résultent des lois correspondant à des catégories d'interactions possibles entre les structures visuelles. Le tableau ci-dessous synthétise ces principales lois qui composent la théorie de la Gestalt [Card *et al.*, 1999b; Chang *et al.*, 2002; Ware, 2000; Zlatoff *et al.*, 2004] :

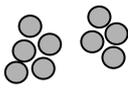
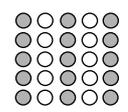
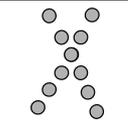
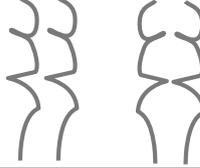
Lois	Effets	Illustrations
<i>Prégnance</i>	Une image est facile à comprendre si sa structure est simple et inversement.	Voir ci-dessous
<i>Proximité</i>	Deux composants qui sont proches ont tendance à être perçus comme un seul composant. Dans l'illustration à droite, l'image est perçue comme deux éléments par proximité de points.	
<i>Similarité</i>	Les composants similaires sont perçus comme s'ils étaient regroupés. Dans l'illustration à droite, l'image est perçue comme une série de colonnes par similarité de couleurs.	
<i>Fermeture</i>	Les contours proches sont perçus comme unifiés. Dans l'illustration à droite, on perçoit un triangle blanc alors qu'il n'est pas dessiné.	
<i>Continuité</i>	Des éléments voisins sont perçus groupés lorsqu'ils possèdent potentiellement un trait qui les relie. Dans l'illustration à droite, l'image est perçue comme une croix.	
<i>Symétrie</i>	Des éléments sont perçus comme un élément global lorsqu'ils forment une symétrie. Dans l'illustration à droite, les deux éléments de gauche ne sont pas perçus comme un seul élément contrairement aux deux suivants.	
<i>Trajectoire identique</i>	Des éléments qui se déplacent avec la même trajectoire semblent groupés.	
<i>Familiarité</i>	Des éléments sont plus facilement groupables si le groupe est familier ou significatif.	

Figure 12 - Principales interactions possibles entre les structures visuelles.

La loi dominante est la loi de *prégnance*. En effet, l'œil privilégiera davantage la reconnaissance de formes simples. Par exemple dans l'image ci-dessous, comment percevez-vous la forme de la zone grise ?

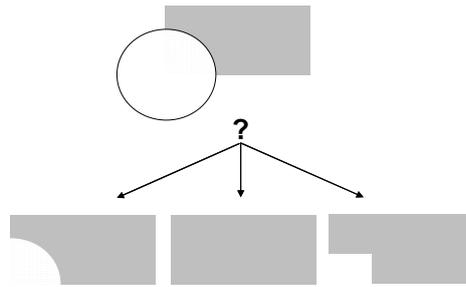


Figure 13 - Loi de profusion : l'œil recherche toujours le motif le plus simple.

Ces lois permettent aussi d'expliquer de nombreuses illusions d'optiques basées sur une perception ambiguë.

Par exemple avec l'image ci-dessous : s'agit-il d'un vase ou de deux visages ?

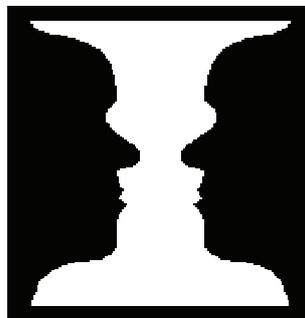


Figure 14 - Perception ambiguë.

d. La carte comme support de la pensée

L'objectif de cette étude est d'appliquer la cartographie à la gestion des connaissances comme paradigme d'accès aux connaissances des organisations.

Prise dans son sens le plus général, la cartographie permet de construire des cartes (support externe graphique) pour aider les hommes à appréhender leurs différents espaces. Elle met en œuvre des techniques permettant de passer d'un espace informationnel à un espace d'information cartographique constituant une carte.

Au fil des siècles, la cartographie est passée du statut d'art graphique à celui d'outil d'analyse, d'aide à la décision et de communication. Elle constitue une science qui place la carte au-delà de la simple image artistique d'un espace informationnel. En effet, la carte est devenue un instrument de transmission de message mais surtout un instrument puissant d'aide à la décision.

Le recours à la cartographie permet d'exploiter et de valoriser deux caractéristiques naturelles de l'homme :

- > **La cognition externe** : nous sommes capables d'utiliser notre environnement pour créer des représentations externes jouant le rôle d'extension de notre mémoire. Ces représentations nous permettent alors d'effectuer des raisonnements distribués et par voie de conséquence, ce mécanisme permet de réduire notre effort cognitif.

- > **La perception d'information graphique** : l'information cartographique contenue dans une carte est graphique. Cette classe de support externe exploite les capacités naturelles de l'homme à percevoir et à traiter l'information graphique.

La cartographie nous offre la possibilité d'exploiter ces deux capacités pour faire des cartes un support formidable de la pensée. C'est pourquoi, elle constitue l'axe principal de notre approche pour résoudre notre problématique de maîtrise de l'espace informationnel des organisations.

2.2 La cartographie de données abstraites

La cartographie de données abstraites appliquée aux organisations a pour finalité de produire des cartes permettant de visualiser leur espace informationnel. Ces informations sont dites abstraites car elles ne sont pas issues de mesures physiques (comme des relevés topographiques indiquant la position des villes d'un pays). Par conséquent, les données abstraites ne possèdent pas de représentation graphique intrinsèque. La difficulté de la cartographie est alors de déterminer le choix des représentations pour les visualiser.

2.2.1 Une activité pluridisciplinaire

La cartographie de données abstraites appliquée à l'espace informationnel des organisations est une activité intrinsèquement pluridisciplinaire.

Tout d'abord, elle doit mettre en œuvre des méthodes et des techniques du domaine de la **gestion des connaissances** pour manipuler les informations des organisations (comme des techniques de modélisation et de structuration d'espace informationnel).

Ensuite, la cartographie a pour finalité de produire des cartes contenant des représentations graphiques (quelle que soit la nature des informations). Face au nombre infini de représentations possibles pour coder une information (combinaison des valeurs possibles pour les variables visuelles comme la forme, la couleur, etc.), la **sémiologie graphique** permet d'étudier la signification de chaque type de représentation et ainsi de guider leur choix.

Pour finir, les **sciences cognitives** permettent de comprendre l'impact d'une carte sur une personne et de mettre ainsi en correspondance le message véhiculé par celle-ci avec le message souhaité par son concepteur.

La cartographie de données abstraites est donc à la croisée de la gestion des connaissances, de la sémiologie graphique et des sciences cognitives :

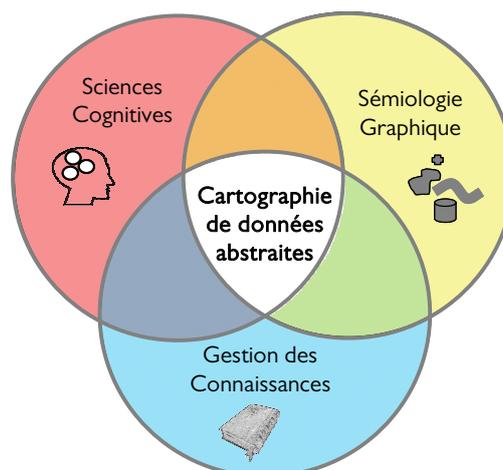


Figure 15 - Cartographie de données abstraites : une activité pluridisciplinaire.

L'étude d'un domaine pluridisciplinaire a pour particularité d'obliger à manipuler des terminologies multiples. Bien que les concepts manipulés soient les mêmes, les mots d'usage

utilisés pour les dénoter peuvent varier. Encore plus délicats sont les mots d'usage communs à plusieurs terminologies mais, qui dénotent des concepts différents dans chaque domaine.

Face à cette problématique, nous nous sommes tout d'abord approprié les terminologies de chaque domaine puis nous avons proposé et utilisé notre propre terminologie. Elle est le fruit de notre expérience ; à la croisée de nombreux domaines, elle en reprend des termes avec pour objectif d'être précis et consensuel. Malheureusement ne pouvant pas être toujours consensuel, les spécialistes de chaque domaine pourront être en désaccord avec l'emploi de certains termes. Le cas échéant, l'important est alors de s'attacher aux concepts manipulés et moins aux termes pour les dénoter.

2.2.2 Les approches courantes

Pour cette étude, la cartographie a pour finalité de produire des cartes permettant de visualiser des espaces informationnels. Elle peut être assimilée à un processus de transformation d'un ensemble d'informations en une visualisation.

Il existe de nombreux travaux qui traitent de cette transformation cependant, ils ne sont pas associés au domaine de la cartographie bien qu'ils traitent de ce thème.

Dans l'esprit de la plupart des auteurs, la cartographie n'est pas dissociable de la géographie. Pourtant son histoire nous montre bien qu'elle a toujours tendu à représenter des informations non géographiques. L'invention des cartes thématiques témoigne bien de cette tendance : le champ des informations traitées par la cartographie s'étend à toutes les informations (et pas seulement les informations géographiques).

Pour parler de cartographie de données non géographiques, la littérature parle davantage de « visualisation ».

Remarque : par définition, la visualisation est plus restrictive que la cartographie car la visualisation est l'« action de rendre visible un phénomène qui ne l'est pas »¹ et la cartographie a pour objet de construire et d'utiliser une carte (qui correspond à une visualisation). La cartographie inclut donc la visualisation ainsi que ses techniques.

Les travaux liés à la visualisation sont associés à deux communautés distinctes. La première, qui est aussi la plus ancienne, est la communauté de la « **visualisation d'informations** ». La deuxième est celle de la « **visualisation de connaissances** ».

Ces deux communautés proposent deux approches différentes pour visualiser un espace informationnel. Pour bien comprendre ce qui les différencie, il est important de bien comprendre ce que nous associons aux concepts transversaux de ces deux approches : donnée, information et connaissance.

a. Donnée, information et connaissance

Dans le domaine de la visualisation (d'information ou de connaissances) et plus généralement dans le cadre de la gestion des connaissances, les termes de « donnée », « information » et « connaissance » sont récurrents. C'est aussi le cas pour cette étude, c'est pourquoi il est important de préciser et de comprendre le sens que nous leur attribuons.

Malheureusement, bien que ces concepts soient manipulés par de nombreuses communautés, nous n'avons pas réussi à leur trouver des définitions précises et consensuelles.

¹ Trésor de la Langue Française. <http://www.tlfi.fr>.

De plus, il n'est pas rare que les auteurs d'une même communauté ne soient d'accord. C'est pourquoi, il nous semble utopique de réussir ici à proposer une définition consensuelle de ces concepts. Cette section a cependant pour objectif de présenter notre vision. Elle nous appartient et nous gardons à l'esprit qu'elle peut diverger de celle d'autres auteurs.

Pour débiter, nous avons étudié les définitions proposées par des ouvrages de référence comme le « Trésor de la Langue Française »¹ et le « Dictionnaire de l'Académie 9^{ème} édition »². Il nous semble que ces définitions constituent un point de départ légitime.

i. Donnée et Information

Pour le Trésor de la Langue Française, la donnée est « ce qui est connu et admis et qui sert de base, à un raisonnement, à un examen ou à une recherche » et l'information est un « élément de connaissance susceptible d'être représenté à l'aide de conventions pour être conservé, traité ou communiqué ». La donnée est donc similaire au fait qui est perçu et l'information est sa représentation que l'on peut manipuler.

Pour le Dictionnaire de l'Académie 9^{ème} édition, c'est l'inverse, la donnée est une « représentation d'une information sous une forme conventionnelle adaptée à son exploitation ».

Face à cette confusion et pour cette étude, nous choisissons de regrouper les concepts de donnée et d'information car, quelles que soient les définitions considérées, la frontière qui les sépare (quand elle existe) nous semble beaucoup trop mince. C'est pourquoi dans le cadre de la cartographie sémantique, nous estimons que les données et les informations sont similaires. Pour les définir, nous reprenons la définition qui est faite de « donnée » dans le Dictionnaire de l'Académie 9^{ème} édition : « Ce qui est connu immédiatement par le sujet, indépendamment de toute élaboration de l'esprit, par opposition à ce qui est connu par induction ou déduction, par raisonnement, par calcul ».

ii. Connaissance

Pour le Dictionnaire de l'Académie 9^{ème} édition, la connaissance est une « idée, notion que l'on a d'une personne ou d'une chose, représentation que l'on s'en fait ». Pour le Trésor de la Langue Française, la connaissance consiste à « établir le modèle ou la description simplifiée d'un phénomène, d'un processus ou d'un système, en vue d'en étudier le fonctionnement par simulation ».

De manière générale, le concept de connaissance est lié au fait de posséder une représentation mentale d'un fait. Pour faire le lien avec les données et les informations, on peut dire qu'il y a connaissance lorsqu'on dispose face à une information (ou donnée) de concepts en relations avec cette information.

Pour conclure sur cet aspect définitoire de donnée, information et connaissance nous proposons la synthèse suivante : la cartographie de données abstraites est assimilable à un processus qui permet de passer de données à une carte où les données peuvent décrire des connaissances.

Maintenant que notre vision (de donnée, information et connaissance) est présentée, nous pouvons reprendre la description des deux approches de la visualisation en commençant par la visualisation d'informations.

¹ <http://www.tffi.fr>

² <http://www.academie-francaise.fr/dictionnaire>

b. La visualisation d'informations

L'approche proposée par la communauté de la visualisation d'informations (souvent aussi nommée « InfoVis » ou « IV ») trouve ses origines dans l'étude des interfaces homme-machine (IHM). Pour cette communauté, la problématique de la visualisation (et par conséquent de la cartographie) peut être résumée à la question suivante :

« Comment représenter un grand nombre d'informations sur un écran ? »

i. Différence avec la visualisation scientifique

En 1999 est paru un recueil d'articles élaboré par S. K. Card, J. D. Mackinlay et B. Shneiderman nommé « Readings in information visualization : using vision to think » [Card et al., 1999d]. Dans ce livre, ces auteurs ont écrit un article d'introduction qui est une synthèse, reconnue par leurs pairs, sur le domaine de la visualisation d'informations [Card et al., 1999b].

Pour commencer, ils définissent la visualisation (en générale) par :

« La visualisation est l'utilisation de représentations visuelles interactives et informatisées de données pour amplifier la cognition ».

La visualisation a donc pour objectif de représenter graphiquement des données pour permettre aux utilisateurs des cartes de mieux raisonner. Il est possible de remarquer l'héritage des IHM avec le caractère nécessairement informatisé de la visualisation.

Ensuite, les auteurs subdivisent le domaine de la visualisation en deux classes selon la nature des données. Dans la première classe, ils regroupent sous le nom de « **visualisation scientifique** » toutes les visualisations de phénomènes physiques. La cartographie géographique entre dans cette classe. Dans la seconde, ils regroupent cette fois sous le nom de « **visualisation d'informations** » toutes les visualisations de données abstraites (non liés à des phénomènes physiques).

Ils définissent la visualisation d'informations comme :

« La visualisation d'informations est l'utilisation de représentations visuelles interactives et informatisées de *données abstraites* pour amplifier la cognition ».

Pour les auteurs, on ne peut pas cartographier de la même manière des données scientifiques et des données abstraites.

Pour comprendre cette discrimination entre visualisation de données physiques et visualisation de données abstraites, il est nécessaire de comprendre la difficulté liée à la représentation de données abstraites par rapport à des données scientifiques.

La visualisation scientifique est un outil qui permet de manipuler un grand nombre de données scientifiques afin de permettre de « voir » des phénomènes issus de ces données. La visualisation scientifique est donc basée sur des données physiques (la terre, des molécules, le corps humain) déjà localisées dans un espace. Pour Card, Mackinlay et Shneiderman tant que les informations et les données que l'on souhaite représenter sont dérivées de données physiques, alors elles possèdent une représentation graphique intrinsèque. C'est le cas par exemple avec une carte géographique qui représente des villes ; la localisation physique des villes sur le globe terrestre fournit la position relative des points correspondants aux villes sur la carte. Les représentations graphiques sont alors « à l'image » des phénomènes observés. Pour les données abstraites, la difficulté est de leur trouver une représentation et plus particulièrement une répartition dans l'espace de la carte.

Pour illustrer cette particularité, supposons que l'on souhaite cartographier un ensemble de concepts. La difficulté n'est pas de trouver une représentation pour chaque concept (par exemple un rectangle ou un cercle) mais bien de les répartir dans l'espace de la carte sachant que la répartition aura un impact important sur la signification perçue.

Il est donc facile de comprendre que les données abstraites n'ont pas de représentation graphique intrinsèque. C'est pourquoi, la principale problématique de la visualisation d'informations est de déterminer quelles représentations choisir pour véhiculer quel message.

La section suivante, présente un exemple de visualisation d'informations.

ii. Exemple de visualisation d'informations

Pour comprendre la difficulté de manipuler de grands ensembles de données abstraites et l'intérêt de passer par une cartographie, prenons un exemple réel : l'étude de l'émergence et de l'évolution d'un domaine scientifique [Börner *et al.*, 2005].

L'objectif de ces travaux était d'effectuer de la veille sur un domaine scientifique donné. Plus particulièrement, il s'agissait pour les auteurs de proposer un système pour étudier l'impact dans le temps de chaque auteur sur le domaine donné.

Pour leur étude, les auteurs ont analysé un ensemble de données qui illustre la naissance et la croissance d'un nouveau domaine des sciences de l'information entre 1974 et 2004 : la visualisation d'informations¹. L'ensemble des données contient tous les articles de la bibliothèque d'ACM² liés à la recherche sur la visualisation d'informations : soit 614 articles écrits par 1036 auteurs uniques entre 1974 et 2004.

Pour réussir à comprendre la structure du domaine et son évolution, les auteurs ont exploité les informations contenues dans le jeu de données ; ils ont basé leur étude principalement sur le nombre de publications pour chaque auteur ainsi que leurs citations. Le jeu de données est structurable en un graphe de coauteurs. Même limité à ce graphe, le nombre d'informations est immense.

Dans ce contexte-là, comment appréhender le domaine étudié ? Comment déterminer l'auteur qui influence le plus le domaine étudié ? Tout le monde a conscience que la réponse est dans le jeu de données et plus particulièrement dans le graphe. Mais comment manipuler ce graphe ? Un traitement simple permet de calculer l'auteur qui a le plus publié ou alors, l'auteur qui est le plus cité ou bien encore, le binôme d'auteurs qui est le plus cité et ainsi de suite.

Face à ce jeu de données, les auteurs ont calculé un ensemble de métriques pour chaque auteur (du jeu de données) en fonction de leur nombre de publications et leur nombre de citations. Voici quelques résultats avec les dix premiers auteurs de chaque métrique : le degré (degré du sommet correspondant à chaque auteur dans le graphe), la force de productivité (nombre de papiers produits), la force de citation (nombre des citations reçues) et la centralité (chemins les plus courts qui traversent cet auteur dans le graphe).

¹ Ce jeu de données a été rendu disponible pour le concours d'InfoVis en 2004. <http://www.cs.umd.edu/hcil/iv04contest>

² <http://portal.acm.org>

Degree k	#	Productivity Strength S_p	#	Citation Strength S_c	#	Betweenness	#
B_Shneiderman	23	B_Shneiderman	7.62	S_K_Card	88	B_Shneiderman	10893
J_D_Mackinlay	17	S_K_Card	5.71	J_D_Mackinlay	67	S_K_Card	10618
S_K_Card	17	J_D_Mackinlay	4.37	B_Shneiderman	66	J_D_Mackinlay	8357
G_Robertson	16	Daniel_A_Keim	4.11	G_Robertson	64	Stephen_G_Eick	7420
Allison_Woodruff	15	Steven F. Roth	3.96	Christopher Ahlberg	36	Chris_Olston	5165
Lucy_T_Nowell	15	John_T_Stasko	3.92	R_Rao	34	Ben_Bederson	4791
Roberto_Tamassia	15	Stephen_G_Eick	3.67	Ben_Bederson	25	Mei_C_Chuah	4718
Ben_Bederson	15	G_Robertson	3.46	Peter_Pirolli	21	G_Robertson	3187
Harpreet_S_Sawhney	14	Ben_Bederson	3.40	Steven_F_Roth	20	Steven_F_Roth	2063
M_Stonebraker	14	Marc_H_Brown	3.33	Brian_Johnson	17	E_H_-H_Chi	1718

Figure 16 - Exemple de cartographie : construction de métriques.

Mais l'influence d'un auteur (ou un binôme) ne peut pas se résumer à la somme de ces quelques mesures. Les auteurs ont donc pondéré chacun de ces résultats pour obtenir un poids global. Une fois cette pondération effectuée pour chaque auteur, il faut ensuite la comparer à celle des autres auteurs.

Une approche « basique » sous forme de table de données permet de répondre à des questions liées à un, voir deux auteurs (**vision locale**). Mais au-delà comment faire ? Comment appréhender tout le domaine pour comprendre l'ensemble des interactions entre les auteurs ? Comment avoir une **vision globale** ?

Pour répondre à leurs besoins et face à la grande quantité d'informations, les auteurs ont choisi de cartographier l'espace informationnel constitué du jeu de données. Ils ont obtenu plusieurs représentations graphiques du graphe.

Ainsi, ils ont obtenu une série de cartes représentant le graphe pour différentes périodes :

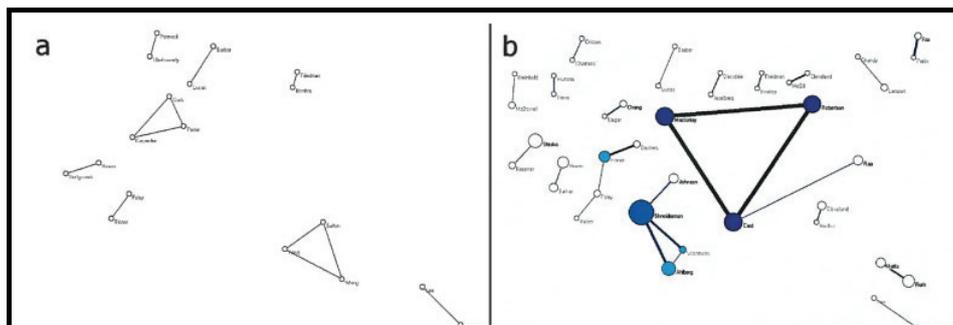


Figure 17 - Exemple de cartographie en visualisation d'informations [Börner *et al.*, 2005] : partie a & b.

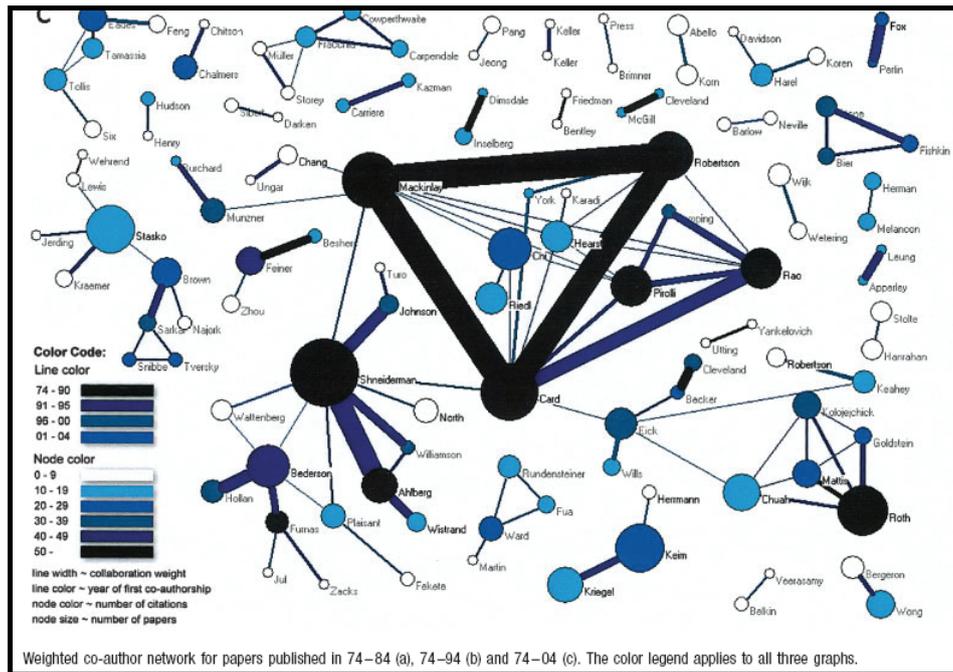


Figure 18 - Exemple de cartographie en visualisation d'informations [Börner *et al.*, 2005] : partie c.

Chaque carte ne contient que les auteurs les plus influents (selon les paramètres de la cartographie) ainsi que leurs relations. Sans même lire la légende associée à la carte, il est possible d'identifier immédiatement les auteurs les plus influents mais aussi les différentes communautés. Dans la dernière carte, on peut lire que les personnes les plus influentes sont celles qui constituent le trio : Mackinlay, Card et Robertson.

Ce petit exemple permet donc bien de comprendre le gain offert par la cartographie de données abstraites en visualisation d'informations face à un grand nombre d'informations.

iii. Modèle de référence de la visualisation d'informations

Tous les auteurs s'accordent à dire que la visualisation d'informations est un processus qui permet de passer de données à une forme graphique. Card, Mackinlay et Shneiderman ont proposé une modélisation plus fine de ce processus qui est devenue rapidement le « modèle de référence » de la visualisation d'informations [Card *et al.*, 1999b].

Voici le schéma de leur modèle ; les flèches représentent des flots de données :

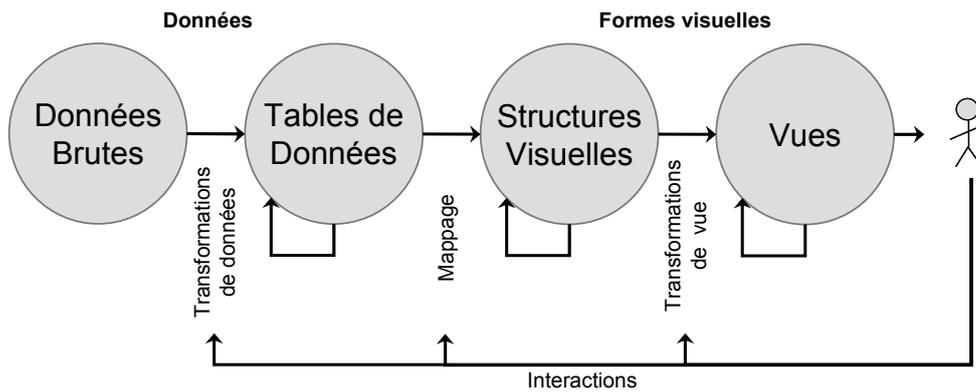


Figure 19 - Modèle de référence de la visualisation d'informations.

Dans ce modèle, la visualisation est décrite comme un processus qui passe par des « états de données » différents. Le premier correspond aux **données brutes** à traiter. Elles sont dans un format propre (par exemple un fichier de calcul ou un document textuel).

L'état suivant est obtenu par une transformation des données brutes pour obtenir une structure **de table de données**. Les tables de données sont définies comme un ensemble de relations représentées sous forme de matrice dont les tuples (colonnes de la matrice) définissent des variables et les valeurs des tuples définissent des cas (lignes de la matrice).

Cette dernière est ensuite transformée en un ensemble de **structures visuelles** par une opération de mappage visuel. Cette opération consiste à associer aux éléments de la table de données, un ou plusieurs objets graphiques (les structures visuelles).

Et enfin, les structures visuelles sont transformées en **vue** par une opération de transformation de vue. Les opérations permettant de passer d'un état de données à un autre ne sont pas uniques ; c'est pourquoi, il est possible de construire un très grand nombre de vues à partir d'un même ensemble de données brutes.

La phase essentielle de ce modèle est la transformation des tables de données en structures visuelles : le mappage. En effet, les tables de données contiennent des données abstraites alors que les structures visuelles sont des éléments graphiques faisant varier des propriétés visuelles perçues par la vision humaine (les variables visuelles). L'opération de mappage est l'opération qui permet de passer d'un monde de données à un monde de formes visuelles.

Par la suite, Ed H. Chi a proposé une taxonomie des techniques visuelles basées sur ce modèle [Chi & Riedl, 1998; Chi, 2000]. Dans ce dessein, il a considéré quatre niveaux dans le modèle. Ces niveaux, il les nomme des « niveaux d'abstraction des données dans le modèle de référence » :

- > Données,
- > Abstraction analytique (correspondant aux tables de données),
- > Abstraction visuelle (correspondant aux structures visuelles) et,
- > Vue.

Il a ensuite décomposé toutes les techniques visuelles connues en une succession d'opérations atomiques permettant de passer des données à la carte. Les opérations permettent alors de passer d'un niveau à un autre (non nécessairement de natures différentes).

Les passages entre des niveaux de natures différentes permettent alors de définir des classes d'opérations correspondant au modèle de référence :

- > Les opérations de transformations des données : elles permettent la sélection et la transformation d'un ensemble de données pour le visualiser. Les données sont stockées dans une table de données ;
- > Les opérations de mappage : passage d'une table de données à un ensemble de structures visuelles ;
- > Les opérations de transformations visuelles : les structures visuelles sont rendues visibles. Ces opérations déterminent la perception de la carte.

Par la suite d'autres travaux ont repris cette approche du processus de visualisation. La littérature parle généralement de « pipeline de la visualisation ». On peut notamment citer les travaux de [dos Santos & Brodlie, 2004] qui introduisent deux étapes intermédiaires dans le modèle de référence : une étape d'analyse suivie d'une étape de sélection des données à cartographier.

La section suivante présente la deuxième approche pour cartographier des données abstraites.

c. La visualisation de connaissances

Tout comme la visualisation d'informations, la visualisation de connaissances a pour fonction de représenter graphiquement des données abstraites.

i. Les origines

La visualisation de connaissances est une communauté plus récente que celle de la visualisation d'informations. On retrouve ses origines dans la gestion des connaissances.

En gestion des connaissances, les connaissances constituent la ressource élémentaire. Donc, selon l'acception que l'on fait de la connaissance, les stratégies mises en œuvre diffèrent [Burkhard, 2005] :

- > Si la connaissance est perçue comme un **objet**, la gestion des connaissances a pour objectif de mettre en œuvre des solutions de stockage.
- > Lorsque la connaissance est perçue comme un **processus**, les stratégies se concentrent sur l'optimisation des processus manipulant les connaissances (comme l'identification, la création et le partage de connaissances).
- > Avec la connaissance perçue comme une **capacité**, les stratégies se concentrent sur la mise en valeur de l'avantage stratégique des connaissances (gestion des compétences et construction d'un capital de connaissances).
- > Si la connaissance est considérée comme une condition d'**accès** à l'information, la gestion des connaissances consiste à mettre en place les méthodes pour identifier, retrouver et accéder à l'information.
- > Enfin, si la connaissance est perçue comme un **état** du savoir et de la compréhension, la gestion des connaissances supporte les individus pour étendre leurs connaissances.

Indépendamment des différentes acceptions sur le sens de la connaissance, la gestion des connaissances est assimilable à un processus dynamique qui a trois objectifs majeurs :

- > Optimiser les processus d'entreprise pour les orienter dans une perspective de gestion de la connaissance.
- > Mettre en œuvre des systèmes pour stocker, identifier, retrouver et accéder aux informations et pour supporter la collaboration des individus.
- > Développer une culture de la connaissance pour motiver les personnes à prévoir, créer et partager leurs connaissances.

Tout comme dans le cadre de la visualisation d'informations, la mise en œuvre d'une stratégie de gestion des connaissances impose donc aux individus de manipuler un grand nombre d'informations. Des supports externes sont encore une fois nécessaires pour aider les individus à acquérir, accéder, stocker, restructurer, communiquer et utiliser des connaissances.

Cette problématique constitue l'origine de la visualisation de connaissances.

ii. Les définitions

Du point de vue de Donald F. Dansereau, le concept de « visualisation de connaissances », pris dans son sens le plus strict, est limité aux aspects d'externalisation de la connaissance d'un individu pour lui-même en utilisant un tracé libre [Dansereau, 2005].

Dans la littérature, l'expression de « visualisation de connaissances » est utilisée pour désigner tout procédé permettant de présenter une structure de connaissances (comme d'un expert à des étudiants) ou encore, comme moyen pour évaluer soi-même des connaissances et aider à la compréhension et à la navigation [Keller & Tergan, 2005].

Pour illustrer cette vision de la visualisation de connaissances, prenons l'exemple des cartes conceptuelles [Novak, 2001]. Une carte conceptuelle est une représentation graphique dans laquelle les concepts sont liés entre eux par des liens pour former un réseau. La carte ci-dessous est une carte conceptuelle associée au concept d'exploration humaine :

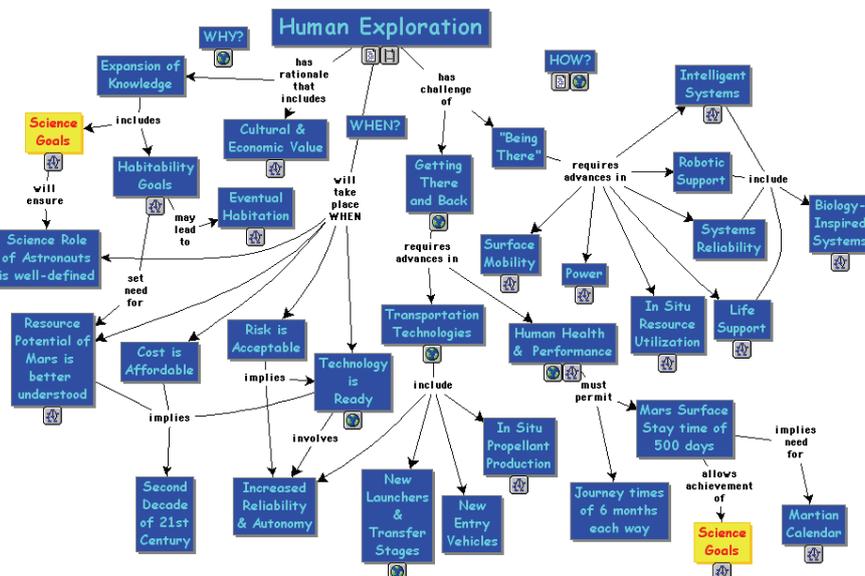


Figure 20 - Carte conceptuelle de l'exploration humaine [Novak, 2001].

Les cartes conceptuelles sont généralement utilisées pour organiser des idées, concevoir et communiquer une structure complexe et résoudre des problèmes.

Jusqu'à maintenant, la visualisation de connaissances porte principalement sur des structures de connaissances conceptuelles (un ensemble de concepts). Cependant, son champ d'application évolue de plus en plus vers d'autres types de connaissances.

Pour Remo Aslak Burkhard, la visualisation de connaissances correspond à « l'utilisation de représentations visuelles pour améliorer le transfert de connaissances entre au moins deux personnes ou groupes de personnes » [Burkhard, 2004; Burkhard, 2005]. D'après cet auteur, les croquis sont une forme courante de visualisation de connaissances car ils servent à représenter rapidement un concept, une idée.

Le croquis ci-dessous, dessiné par Léonard de Vinci, est donc une visualisation de connaissances :

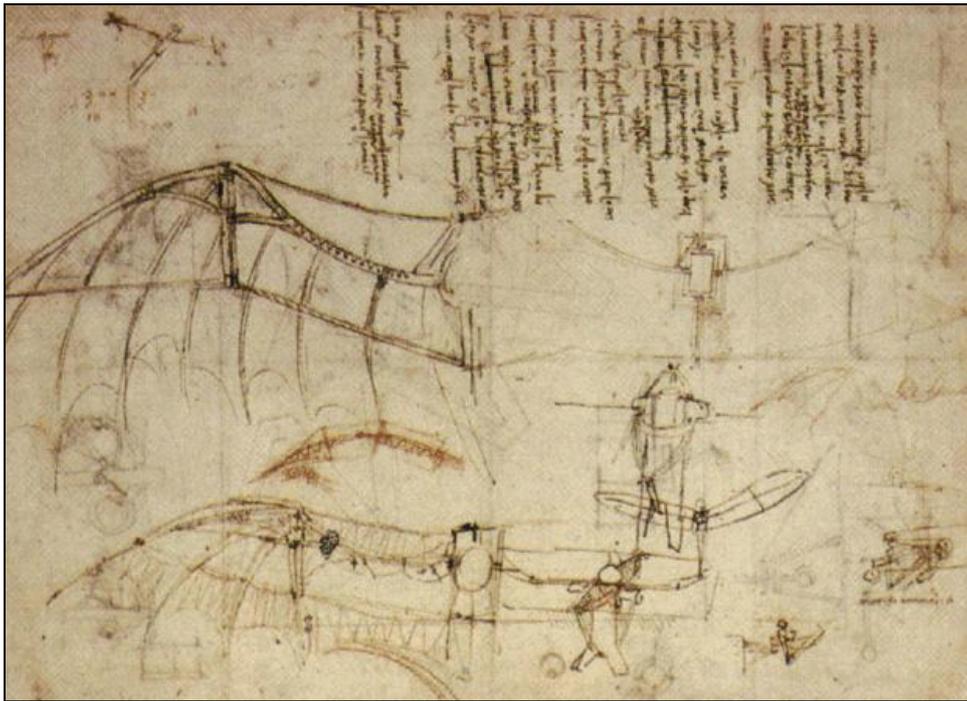


Figure 21 - Croquis de Léonard de Vinci : représentation de l'idée principale du concept de machine volante.

Toujours d'après Remo Aslak Burkhard, l'activité de visualisation de connaissances a tout à gagner à s'inspirer des architectes car ils sont « des experts qui communiquent avec des spécialistes de domaines différents » [Burkhard, 2004]. Pour y parvenir, ils utilisent des représentations complémentaires (des croquis, des diagrammes, des images, des objets et des visualisations interactives) pour amplifier leur cognition et pour communiquer avec les différents corps de métiers qui interviennent sur un chantier de construction.

iii. Différence avec la visualisation d'informations

La littérature qui compare la visualisation d'informations et la visualisation de connaissances est de plus en plus abondante [Keller & Tergan, 2005]. Certains définissent ces deux domaines comme complémentaires et d'autres les définissent comme opposés.

Pour [Jaeschke et al., 2005], la visualisation d'informations porte sur des données abstraites portant sur des objets et les relations entre ces objets, alors que la visualisation de connaissances porte sur des données abstraites qui représentent des connaissances. Pour Remo Aslak Burkard, la visualisation d'informations et la visualisation de connaissances

différent de buts, d'avantages, de contenu, etc. [Burkhard, 2005]. Par exemple, dans sa définition la plus communément admise, la visualisation d'informations se limite à la création de cartes interactives informatisées alors que la visualisation de connaissances ne semble se limiter à aucun média (écran, papier, etc.).

Ces deux visions de la visualisation semblent donc diverger sur l'objet de la visualisation (données abstraites représentant des connaissances ou non). Cependant, leur principale différence est celle des origines : la visualisation d'informations est issue des interfaces homme-machine alors que la visualisation de connaissances est le fruit de la gestion des connaissances. Cette différence d'origine implique malheureusement une différence de communauté, une différence de formation et de parcours pour ceux qui composent ces communautés. Par conséquent, il est difficile de trouver un consensus sur leurs définitions et sur leurs différences.

Selon notre point de vue :

La visualisation d'informations est l'utilisation de représentations visuelles de données abstraites pour amplifier la cognition.

Et,

La visualisation de connaissances est l'utilisation de représentations visuelles de connaissances à des fins de création et de partage.

Les connaissances étant représentées sous forme de données abstraites, la visualisation de connaissances exploite les techniques de la visualisation d'informations. Ainsi, la visualisation de connaissances peut être perçue comme une activité englobant la visualisation d'informations à des fins de gestion de connaissances. Quoi qu'il en soit, ces deux communautés partagent les mêmes paradigmes.

En s'appuyant sur cette analyse, la section suivante présente notre analyse du processus de cartographie de données abstraites en fédérant les deux approches.

2.2.3 Un processus en niveaux

L'histoire de la cartographie nous apprend que quelle que soit la nature des données à cartographier, il s'agit toujours de visualiser un espace informationnel. Mais le passage d'un ensemble d'information à une carte n'est pas immédiat. L'homme ne peut pas percevoir directement les informations de l'espace informationnel. Il passe nécessairement par des *niveaux intermédiaires* [Wattenberg & Fisher, 2003].

Avant d'expliquer cette notion de niveaux, prenons un exemple simple :

le nombre « 7 ».

Ce nombre est inscrit sur le papier que vous lisez et par conséquent vous pouvez le percevoir. Peut-être pensez-vous que le simple fait qu'il soit inscrit sur cette page, suffit à votre cerveau pour accéder directement au nombre 7 ?

Pourtant pour y parvenir, vos yeux perçoivent un graphe ou dessin sur le support externe, la page. Puis un mécanisme, lié à la vision, va identifier une forme constituée de deux segments. Ensuite, le cerveau reconnaîtra cette forme comme étant le symbole associé au chiffre sept et vous prenez enfin conscience du nombre sept.

Si on parcourt le même chemin mais en sens inverse, pour visualiser le nombre sept il est tout d'abord nécessaire de le représenter par deux segments constituant une structure

visuelle. Ensuite, cette structure visuelle doit être visualisée sur le papier par une impression pour le rendre perceptible à l'œil.

D'un point de vue plus général, la construction d'une carte à partir d'un espace informationnel n'est pas immédiate. En effet, il est tout d'abord nécessaire de représenter¹ les entités de l'espace informationnel sous forme graphique avant de les visualiser². Une fois la carte construite, l'utilisateur peut interagir avec elle pour l'adapter à ses besoins.

Dans la cartographie de données abstraites, nous distinguons donc trois niveaux :

1. **Niveau espace informationnel « brut »** : il s'agit de l'espace initial que l'on souhaite cartographier.
2. **Niveau espace informationnel « représenté »** : ce niveau est obtenu en associant des structures visuelles aux éléments de l'espace *brut*. Si l'on reprend l'exemple du nombre sept, sa représentation peut être le symbole « 7 » ou bien les symboles « VII » en numérotation romaine ou bien encore, une barre rectangulaire avec une taille de sept dans un graphique. Pour un espace informationnel donné, il est possible d'en construire une infinité de représentations. Le passage d'un espace à une représentation est l'action de « représenter ». Pour des raisons de simplicité, on qualifie aussi ce niveau de représentation.
3. **Niveau espace informationnel « visualisé »** : ce niveau correspond à la carte. Il est obtenu par une opération de visualisation de l'espace informationnel *représenté*. Le passage d'une représentation à une visualisation est l'action de « visualiser ». Par exemple, le fait d'écrire un texte ou d'imprimer un graphique correspond bien à l'action de visualiser une représentation. Tout comme les niveaux précédents, il est possible de construire plusieurs visualisations pour une même représentation. Pour des raisons de simplicité, on qualifie aussi ce niveau de visualisation.

Remarque : le niveau de l'espace informationnel *visualisé* est le seul niveau « perceptible ». Par conséquent, il n'est pas possible pour l'homme de manipuler un espace informationnel sans représentation et sans visualisation de cet espace.

Avec l'essor de l'informatique, les cartes sont devenues dynamiques et il est à présent possible d'interagir avec elles. Cette évolution fait de l'utilisateur de la carte un « percepteur-acteur » qui navigue dans un espace informationnel. Les interactions sont assimilables à des opérations qui portent sur les différents niveaux de la cartographie.

Par exemple, lorsque l'utilisateur « déplace la carte », le résultat de cette interaction consiste à changer uniquement la visualisation (plus précisément à changer le contrôle de point de vue). En revanche, si l'utilisateur renomme un élément de la carte, cette interaction modifie un élément de l'espace informationnel et entraîne la mise à jour de sa représentation et de sa visualisation.

Voici un schéma qui décrit ce processus de cartographie de données abstraites en niveaux :

¹ Représenter signifie « rendre présent sous la forme d'un substitut, en recourant à un artifice » (Trésor de la Langue Française).

² Le sens de « visualiser » ne correspond pas à l'action de voir – percevoir par le sens de la vue - mais bien à « rendre sensible à la vue ce qui n'est pas naturellement visible » (Trésor de la Langue Française).

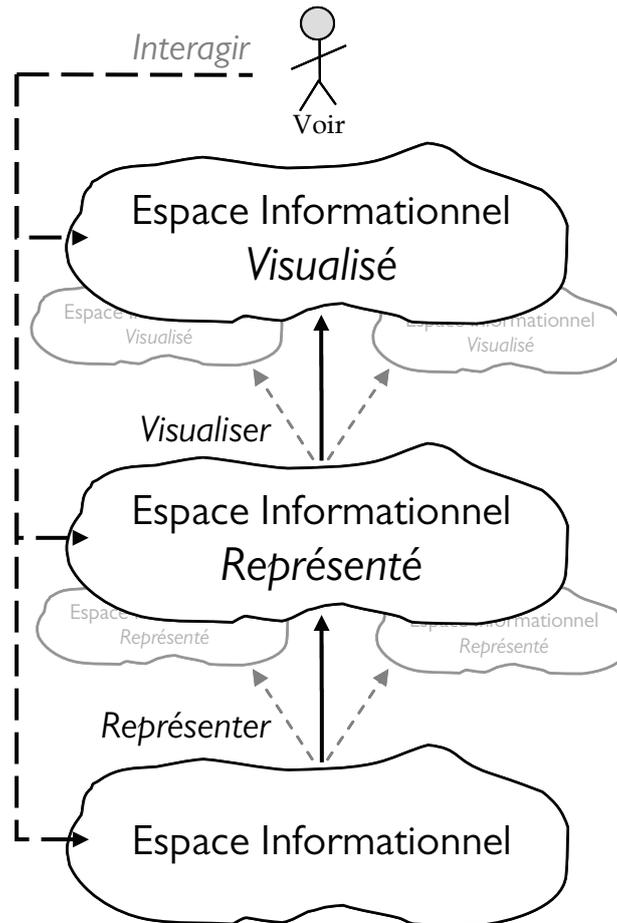


Figure 22 - Les niveaux de la cartographie de données abstraites : de l'information à sa visualisation.

Selon cette approche en niveaux, la cartographie d'un espace informationnel est assimilable à l'application d'une succession d'opérations permettant de passer d'un niveau à l'autre. Le nombre de ces opérations n'étant pas limité à priori tout comme le nombre de leur combinaison, le nombre de cartes qu'il est possible de créer est alors quasiment infini.

Dans ce contexte-là, la problématique est double : comment mesurer l'effet de chaque opération et leur combinaison sur la carte finale ? Comment déterminer les opérations à appliquer pour obtenir le résultat souhaité ?

Cette étude a pour objectif de résoudre cette problématique et par conséquent, pour y parvenir nous devons présenter au préalable les principaux paradigmes qui permettent de construire les différents niveaux nécessaires à la production d'une carte.

2.2.4 Les principaux paradigmes¹

L'objectif de cette section est de présenter les différentes techniques qui permettent de traiter les éléments de l'espace informationnel des organisations pour les visualiser sous la forme de cartes.

¹ À prendre au sens épistémologique du terme : « Conception théorique dominante ayant cours à une certaine époque dans une communauté scientifique donnée, qui fonde les types d'explication envisageables, et les types de faits à découvrir dans une science donnée. » (Trésor de la Langue Française)

Les travaux dans le domaine de la visualisation et plus particulièrement dans celui de la visualisation d'informations, ont permis de produire un grand nombre de paradigmes de cartographie de données abstraites. Chaque paradigme met en œuvre une ou plusieurs techniques et une fois composées, elles permettent de construire différentes cartes d'un même espace informationnel.

Pour comprendre et organiser ces paradigmes, plusieurs auteurs proposent leur taxonomie [Jaeschke *et al.*, 2005; Keim, 2002] dont les principales sont les suivantes :

- > selon le « type des données » à représenter et selon les « tâches » possibles de l'utilisateur [Shneiderman, 1996],
- > selon les structures visuelles utilisées [Lohse *et al.*, 1994],
- > selon les visualisations mises en œuvre [Keim, 2002; Wiss & Carr, 1998],
- > selon les interactions [Chuah & Roth, 1996],
- > selon des transformations de vue [Card *et al.*, 1999d],
- > selon la coordination de vues multiples [Baldonado *et al.*, 2000],
- > selon les aspects cognitifs [Wiss & Carr, 1998] et enfin,
- > selon les étapes du processus de visualisation d'informations [Chi, 2000].

Pour présenter les paradigmes liés à notre problématique, nous n'utilisons aucune des taxonomies présentées précédemment car elles ne correspondent pas à notre approche en niveaux de la cartographie de données abstraites. En effet, dans la plupart de ces taxonomies des paradigmes intervenants à des niveaux différents sont regroupés dans la même classe et inversement.

Par conséquent, pour les présenter tout en restant cohérent avec notre approche, nous utilisons une première classification basée sur le type des paradigmes :

- > Les paradigmes pour représenter des informations (pour obtenir un espace informationnel *représenté*),
- > Les paradigmes pour visualiser ces représentations (pour obtenir un espace informationnel *visualisé*, une carte) et,
- > Les paradigmes d'interactions pour que l'utilisateur interagisse avec la carte.

Ces classes de paradigmes sont décrites dans les sections suivantes.

a. Les paradigmes de représentation

Les paradigmes de représentation permettent de construire un nouvel espace informationnel composé d'un ensemble de structures visuelles localisées. Nous nommons ce nouvel espace un « espace informationnel *représenté* » et parfois plus simplement, une « représentation ».

Les représentations sont obtenues en associant une ou plusieurs structures visuelles aux différents éléments de l'espace informationnel. Chaque structure possède des caractéristiques graphiques correspondant à des variations des différentes variables visuelles (forme, position, taille, etc.). La combinaison des valeurs possibles pour ces variables rend le nombre de représentations possibles quasiment infini.

Les paradigmes de représentation sont très nombreux c'est pourquoi nous avons choisi de les classer selon le type de structure de l'espace informationnel à cartographier : structure « orientée valeurs » ou structure « orientée relations ».

Nous qualifions une structure d'espace informationnel d' « orientée valeurs » lorsque la modélisation qui a permis de la construire se définit comme un ensemble d'entités auxquelles on associe des valeurs. Parfois, certains auteurs parlent aussi de structures vectorielles. Ce type de structure est très courant dans les activités de datamining.

Nous qualifions une structure d' « orientée relations » lorsque la modélisation qui a permis de la construire se définit comme un ensemble d'entités liées par des relations binaires. Ce type de structure est très courant en visualisation de connaissances.

Nous distinguons donc deux classes de paradigmes de représentation. Néanmoins, à la vue des nombreux travaux qui portent uniquement sur la représentation des structures arborescentes (appartenant aux structures orientées relations), nous avons pris le parti de les regrouper dans une classe à part, bien que les arborescences ne soient que des graphes particuliers.

Finalement, pour présenter les paradigmes de représentation nous utilisons trois classes :

- > Les représentations d'espaces informationnels avec une structure orientée valeurs,
- > Les représentations d'espaces informationnels avec une structure orientée relations et,
- > Les représentations d'espaces informationnels avec une structure arborescente.

Chacune de ces classes et leurs principaux paradigmes sont décrits dans les sections suivantes.

Remarque préalable : dans les sections suivantes, les paradigmes sont généralement illustrés par des saisies d'écrans. Étant donné que ces illustrations sont imprimées sur le document, elles sont visualisées (rendu visible). Toute représentation n'est perceptible qu'au travers de sa visualisation par conséquent, toutes les illustrations des paradigmes de représentation illustrent aussi certains paradigmes de visualisation.

i. Représentations orientées valeurs

Les espaces informationnels avec une structure orientée valeurs sont couramment représentés sous forme de table de données (ou matrice). Les entités sont alors nommées des « cas » et elles sont placées en lignes. Les colonnes de ces tables de données sont nommées des « variables ». À l'intersection des cas et des variables, on trouve la valeur associée [Bertin, 1977]. Pour des raisons pratiques, certains auteurs conseillent d'utiliser les tables de données en inversant les lignes et les colonnes [Card *et al.*, 1999b].

Voici une illustration d'une table de données :

	Cas _i	Cas _j	Cas _k	
Variable _x	Valeur _{ix}	Valeur _{jx}	Valeur _{kx}	...
Variable _y	Valeur _{iy}	Valeur _{iy}	Valeur _{ky}	...

Figure 23 - Les tables de données selon [Card *et al.*, 1999b].

Le choix d'un paradigme de représentation est guidé par la dimension de la structure. Dans le cas des structures orientées valeurs, la dimensionnalité correspond aux nombres de variables dans la table de données.

La classe des représentations d'espace avec une structure orientée valeurs est traditionnellement subdivisée en quatre sous-classes selon le nombre de dimensions : une dimension, deux dimensions, trois dimensions et multidimensionnelles.

Pour des raisons de similarité de techniques, nous regroupons les trois premières classes en une seule. Les sections suivantes décrivent ces différentes classes.

Une, deux et trois dimensions

De une à trois dimensions, les espaces sont généralement représentés avec la même technique : les entités (les cas) sont représentées par une structure visuelle dont la position dans l'espace de la carte est fonction des valeurs de ses variables. Dans la table de données associée, les valeurs fournissent les coordonnées de la structure visuelle selon les axes de la carte à raison d'un axe par dimension.

Remarque : il est possible de réduire le nombre de « dimensions graphiques » en codant certaines variables, non pas avec la variable visuelle de position mais avec d'autres variables comme la forme ou la couleur.

Les structures à une dimension correspondent aux listes (par exemple une liste de mots). Dans ce cas, la carte possède un axe. Pour afficher un plus grand nombre d'informations et pour occuper au mieux l'espace de la carte, certaines représentations découpent l'axe en plusieurs parties successives placées en parallèles.

C'est le cas par exemple dans le système Seesoft [Eick *et al.*, 1992]. Il permet de visualiser un listing de code informatique. À chaque ligne de code est associée une structure visuelle sous forme de rectangle de taille proportionnelle et ensuite, elles sont placées sur un axe en fonction de leur position dans le fichier.

L'image ci-dessous est une saisie d'écran de ce système :

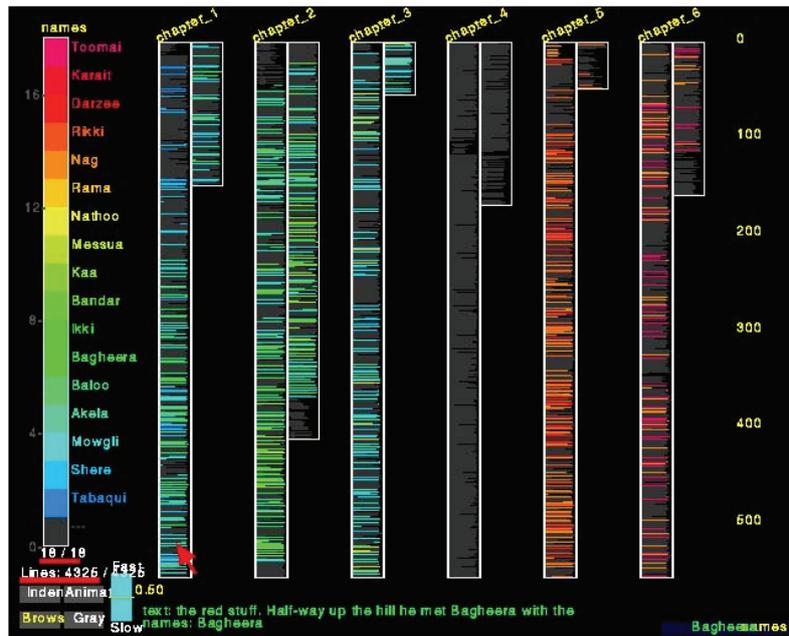


Figure 24 - Seesoft : représentation d'espaces avec une structure à une dimension [Eick *et al.*, 1992].

L'image ci-après correspond quant à elle, à une représentation d'un espace avec une structure à deux dimensions issue du système « SpotFire » [Ahlberg, 1996]. La technique mise en œuvre est nommée « nuages de points » car chaque élément de l'espace (une table de données) est représenté par un point aux coordonnées indiquées par les variables.

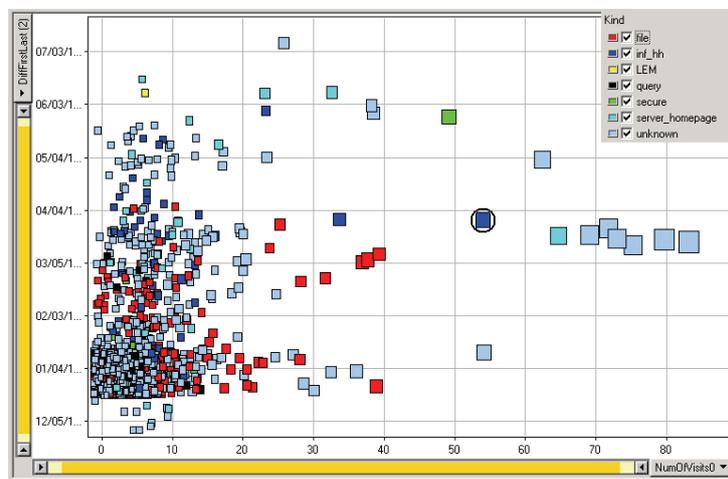


Figure 25 - SpotFire : représentation d'espaces avec une structure à deux dimensions en nuages de points [Ahlberg, 1996].

Pour trois dimensions, il est possible de rajouter une dimension à la carte. C'est le cas avec le système « Themescapes » correspondant à l'image suivante. Ce système permet de cartographier des « clusters » de documents.

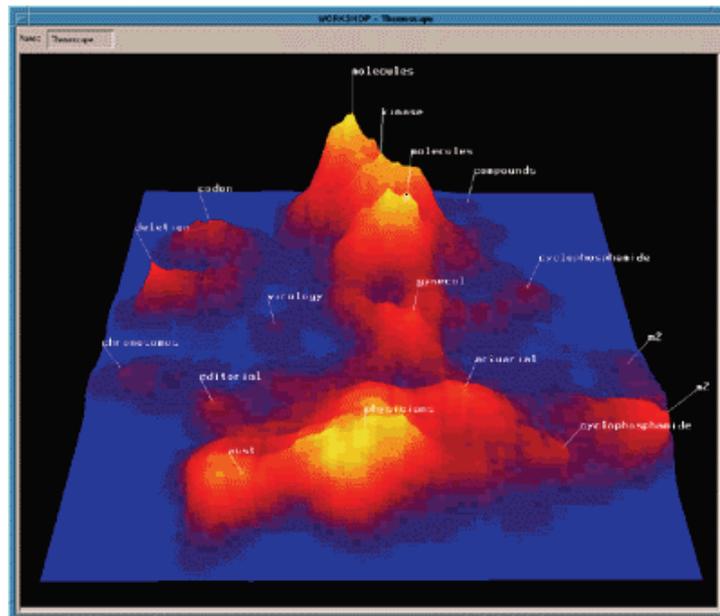


Figure 26 - Themescapes : représentation d'espaces avec une structure à trois dimensions [Wise *et al.*, 1995].

Ces techniques de représentations reposent sur un principe simple : chaque variable de la table de données indique une coordonnée pour placer les structures visuelles sur l'axe associé. On comprend donc la difficulté au-delà de trois dimensions. Néanmoins, il est possible de réduire le nombre de « dimensions graphiques » en codant certaines variables, non pas avec la variable visuelle de position mais avec d'autres variables comme la forme ou la couleur. C'est d'ailleurs souvent le cas, lorsque les valeurs possibles des variables n'appartiennent pas à un ensemble ordonné continu (par exemple, des valeurs qualitatives).

Si l'on reprend l'exemple du système SpotFire présenté précédemment, les deux axes de la carte permettent de coder deux variables. Mais on peut remarquer aussi, que les représentations des éléments possèdent une couleur qui code une troisième variable : le type de cluster.

Au-delà de trois dimensions

Au-delà de trois dimensions, l'espace ne permet pas de créer des cartes avec le même principe que pour des structures de dimensionnalité inférieure. En effet, il n'est pas possible de représenter un espace à plus de trois dimensions avec des axes orthogonaux.

Pour contourner ce problème, une solution consiste à réduire le nombre de dimensions graphique (comme décrit à la fin de la section précédente). Pour cela, il suffit d'utiliser d'autres variables visuelles que la position pour représenter les variables de la table de données.

Il existe aussi d'autres solutions pour représenter des espaces à plus de trois dimensions et les principales sont :

- > Les représentations sous forme de tables de données et quelques variantes
- > Les coordonnées parallèles
- > Les diagrammes en étoile
- > Les figures de Chernoff

Ces quelques techniques sont présentées dans les sections suivantes.

Les tables de données et quelques variantes

Face à un espace informationnel multidimensionnel, l'approche la plus naturelle est de le représenter sous la forme de table de données. Tous les tableurs des suites bureautiques utilisent cette technique.

Le paradigme des « Table Lens » (ou tables optiques) [Rao & Card, 1994] perfectionne cette technique en proposant des améliorations comme la possibilité de représenter chaque valeur (placée à l'intersection entre la variable et le cas) par une structure visuelle plutôt qu'une forme textuelle.

La saisie d'écran suivante provient de ce système :

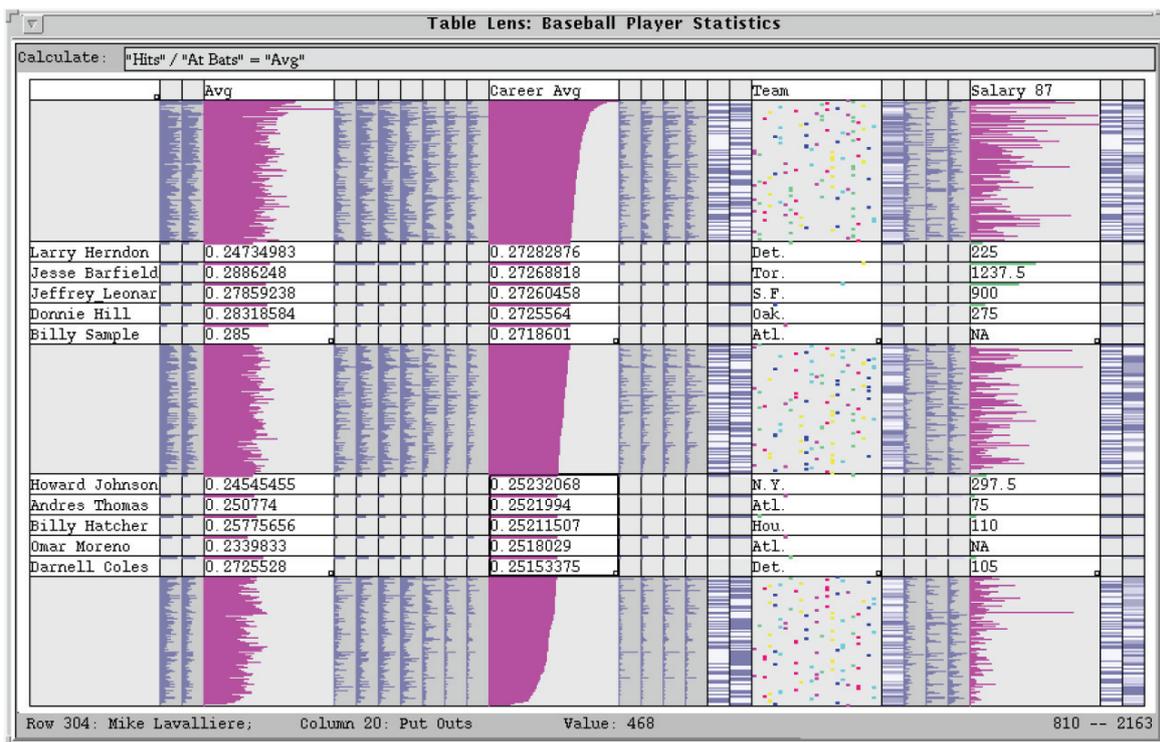


Figure 27 - Table Lens : représentation « Focus+Context » d'une table de données [Rao & Card, 1994].

De plus, le paradigme des tables optiques permet d'appliquer des déformations pour permettre de magnifier certaines plages de cellules et ainsi, les mettre en valeur. Cette technique de déformation est décrite ultérieurement (voir la section sur « Les visualisations non uniformes : focalisation et contexte »).

Les coordonnées parallèles

La technique des « coordonnées parallèles » est une autre technique utilisée pour représenter des espaces multidimensionnels [Inselberg & Dimsdale, 1990]. Les dimensions de l'espace sont représentées par autant d'axes verticaux placés parallèlement et les valeurs des variables sont reportées sur ces axes.

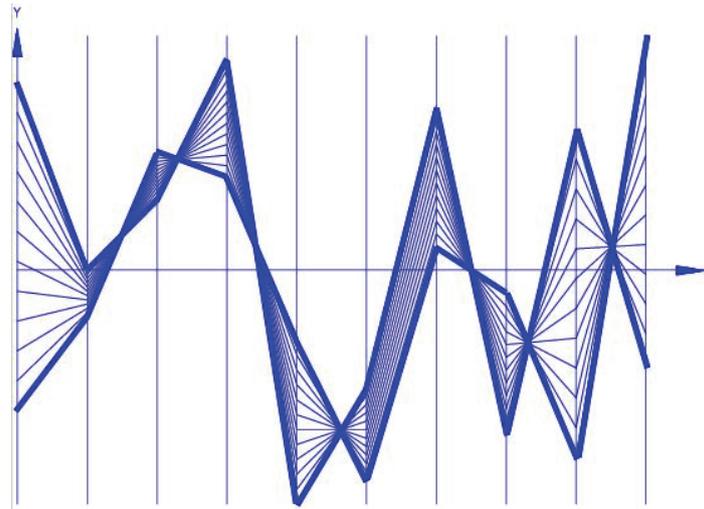


Figure 28 - Cordonnées parallèles : représentation d'espaces avec une structure multidimensionnelle [Inselberg & Dimsdale, 1990].

Les valeurs appartenant au même cas sont reliées par des segments et constituent la représentation d'un cas. L'apparition de faisceaux de trajectoires permet de découvrir des structures sous-jacentes dans l'espace informationnel. Cependant, avec ce type de représentation, la difficulté pour les utilisateurs est d'appréhender chaque cas en détail indépendamment des autres.

Les diagrammes en étoile

Les diagrammes en étoile sont similaires aux coordonnées parallèles car ils permettent de placer de nombreux axes dans la même carte mais avec les axes répartis en « étoile » (plutôt que parallèlement) [Chambers *et al.*, 1983].

L'illustration ci-dessous représente une comparaison de différentes entreprises (les cas) selon différents critères (les variables) comme le profit et le nombre d'employés :

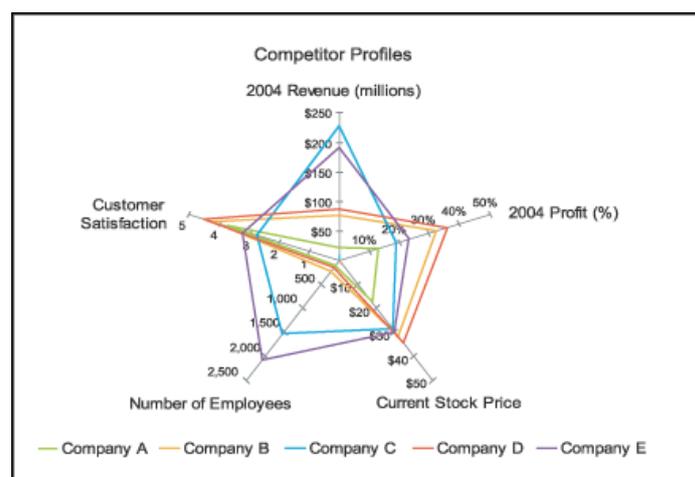


Figure 29 - Les diagrammes en étoile : représentation d'espaces avec une structure multidimensionnelle [Chambers *et al.*, 1983].

Les cas sont là aussi représentés par des segments reliant les différents points sur les axes.

Les figures de Chernoff

Le dernier paradigme de représentation d'espaces multidimensionnels présenté ici se nomme les « figures de Chernoff » [Chernoff, 1973]. Il exploite l'habitude et la forte capacité de l'homme à percevoir de très légers changements dans les expressions faciales.

Une figure de Chernoff s'obtient en associant à chaque variable de la table de donnée un trait d'une expression faciale. Par exemple, la forme de la figure représente la première variable, la taille des yeux représente la deuxième et ainsi de suite pour chaque variable. En utilisant ainsi les traits les plus frappants d'un visage, il est possible de représenter un nombre de dimensions largement supérieur à trois.

L'illustration ci-dessous présente quelques figures de Chernoff :

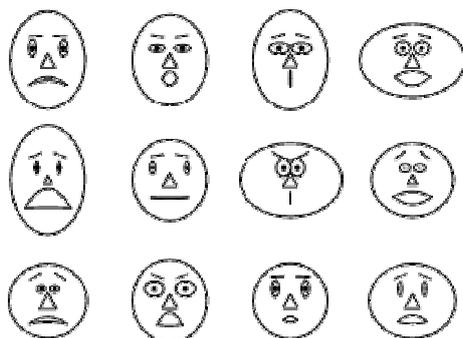


Figure 30 - Les figures de Chernoff : représentation d'espace avec une structure multidimensionnelle [Chernoff, 1973].

ii. Représentations orientées relations

Généralement pour parler de la représentation d'un espace informationnel avec une structure orientée relations, on parle alors de « dessin de graphes ». Il s'agit d'ailleurs d'un domaine de recherche très actif et très prolifique dont les résultats sont mis en œuvre dans de nombreuses activités industrielles (par exemple l'électronique pour dessiner des circuits imprimés).

Pour dessiner un graphe (espace avec une structure orientée relations) il existe deux approches très différentes :

- > Les représentations « nœud-lien » : les entités de l'espace sont représentées par une structure visuelle simple comme un cercle et les tuples des relations sont représentés par des arcs reliant les entités (éventuellement orientés étiquetés).
- > Les représentations matricielles : le graphe est représenté par sa matrice d'adjacence. Il s'agit d'une matrice de booléens (ou de valeurs s'il y a plusieurs relations) où les sommets du graphe (les entités) sont disposés en lignes et en colonnes.

Ces deux types de représentations sont présentés dans les deux sections suivantes.

Représentations « nœud-lien »

Il existe de très nombreux travaux sur les dessins de graphes et plus particulièrement sur les algorithmes de dessin de graphes de type « nœud-lien »¹. Cette problématique fait d'ailleurs l'objet unique de nombreux congrès.

Les représentations nœud-lien (parfois aussi nommés diagrammes), sont utilisées depuis longtemps pour représenter des graphes. Dans ce type de représentations, les entités de l'espace sont représentées par des structures visuelles correspondant aux nœuds du graphe (souvent des cercles) et les tuples des relations sont représentés par des arcs (comme des segments, des flèches, etc.) [Dansereau, 2005].

Il existe de nombreuses techniques de tracés pour représenter les graphes. Le schéma ci-après illustre quelques exemples de tracés parmi les plus répandus [Nishizeki & Rahman, 2004] :

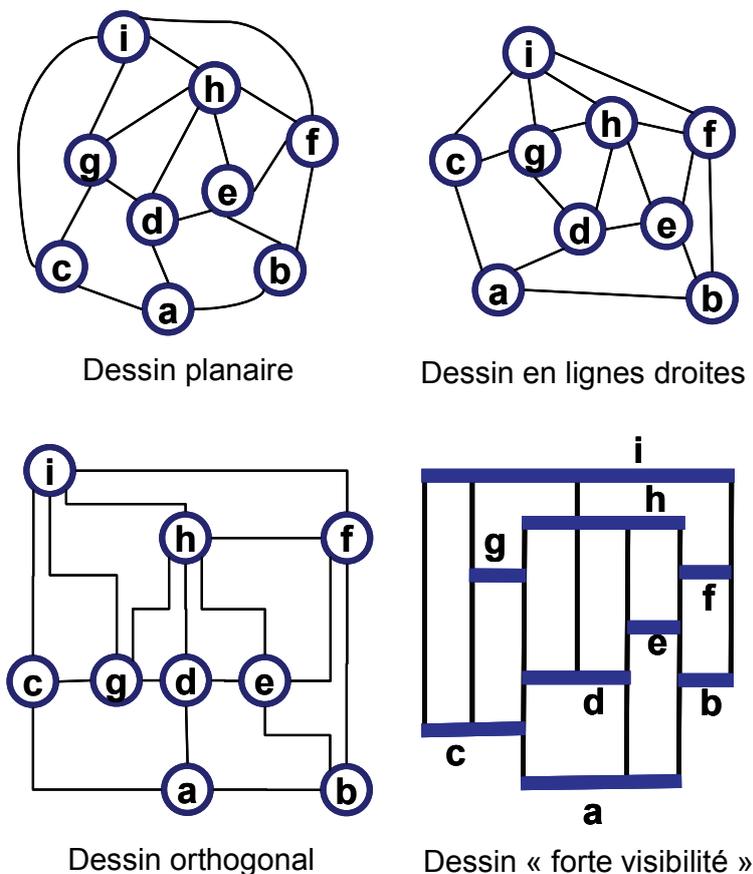


Figure 31 - Exemples de tracés de graphes de type « nœud-lien » extraits de [Nishizeki & Rahman, 2004].

Les représentations nœud-lien sont à l'origine de nombreux travaux et plus particulièrement, de travaux portant sur des techniques de placement des nœuds pour optimiser l'esthétique de la carte (comme la minimisation du nombre d'intersections, la minimisation du rapport entre le lien le plus long et le lien le plus court et la révélation des symétries).

¹ <http://www.graphdrawing.org>

Les algorithmes issus de ces travaux sur le placement des entités et de leurs liens ne sont pas présentés dans ce document car il ne s'agit pas de l'objet de notre étude. Néanmoins, voici quelques exemples de dessins de graphes :

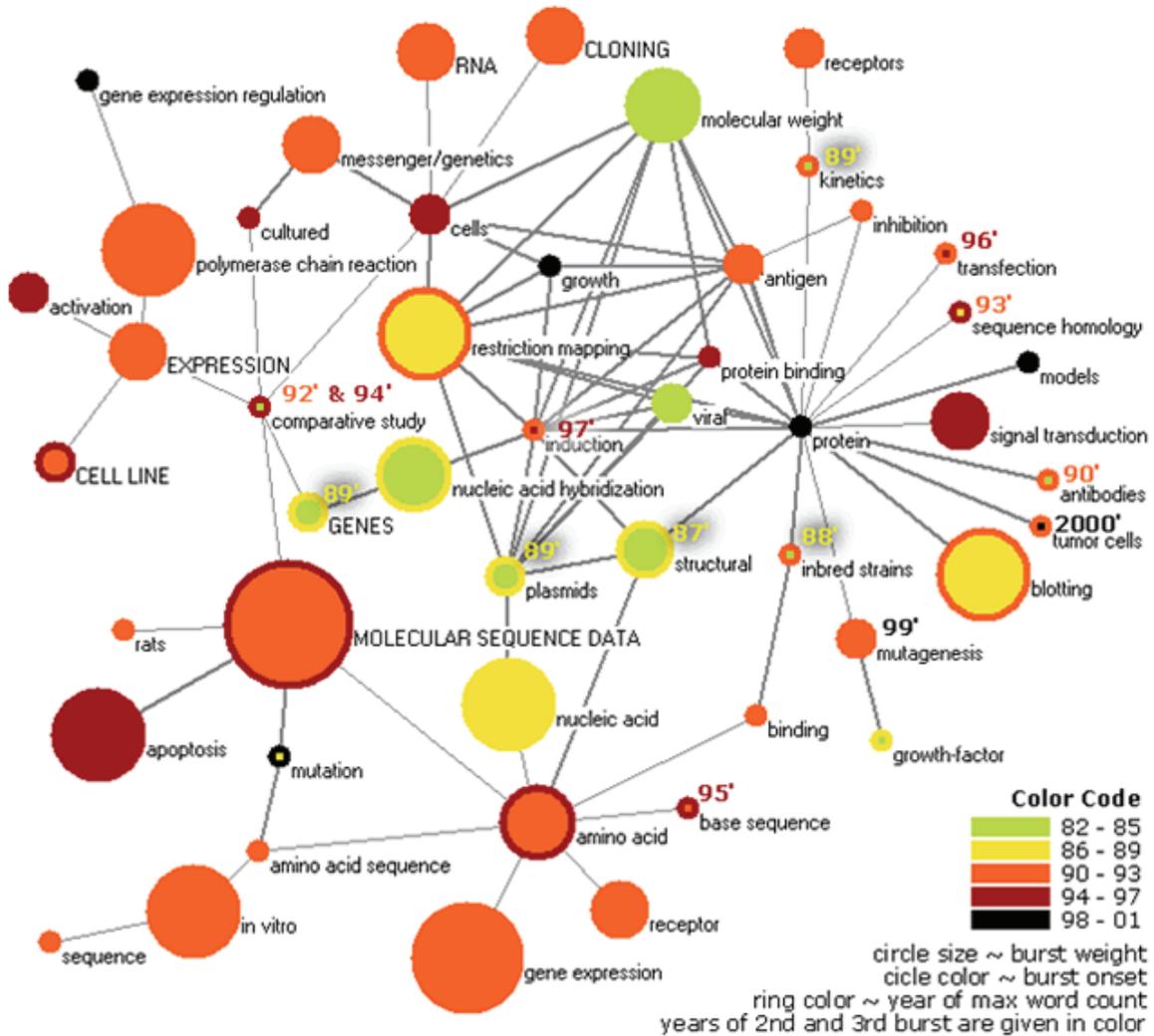


Figure 32 - Représentation d'un graphe des cinquante mots les plus fréquents parmi les 10 % des publications les plus citées des PNAS (« Proceedings of the National Academy of Science ») [Mane & Borner, 2004].

Cet exemple de représentation est intéressant car les structures visuelles associées aux entités de l'espace informationnel (ici des mots) exploitent plusieurs variables visuelles pour coder des informations (ici par exemple, l'importance des mots est codée avec la couleur des disques).

Voici un autre exemple de représentation avec une cartographie de réseaux sociaux :

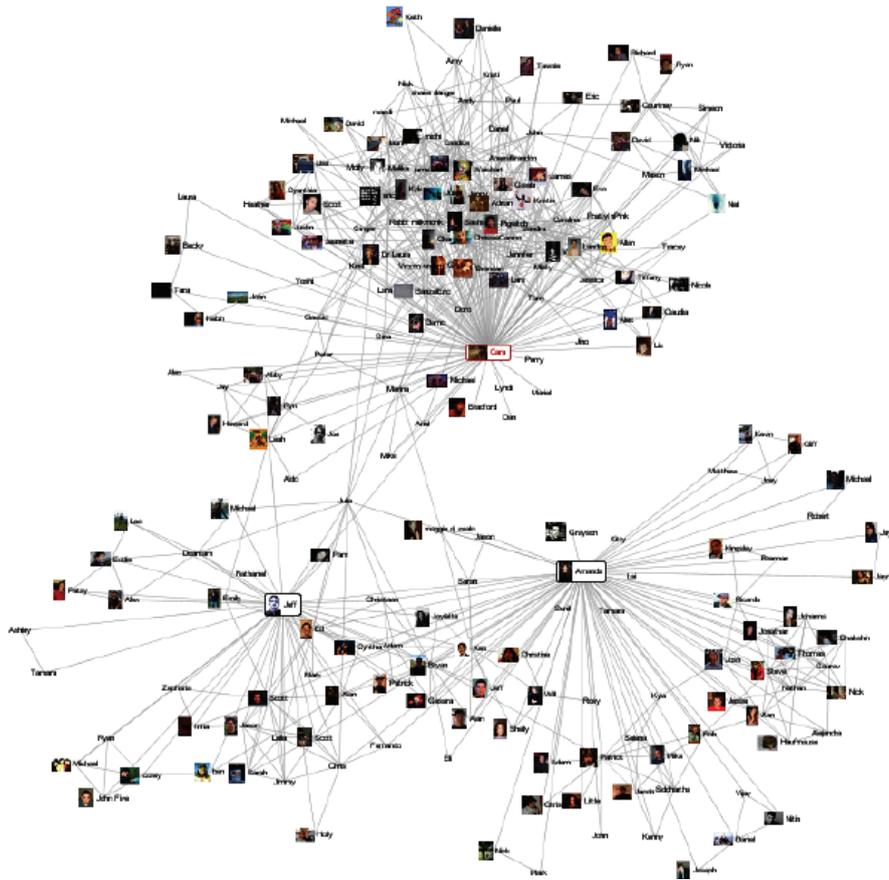


Figure 33 - Représentation de réseaux sociaux [Heer & Boyd, 2005].

Cet autre exemple illustre l'avantage des représentations nœud-lien pour révéler des structures sous-jacentes dans les données et plus particulièrement les composantes connexes et fortement connexes.

D'autres travaux ont eu pour objet de représenter les graphes en trois dimensions :

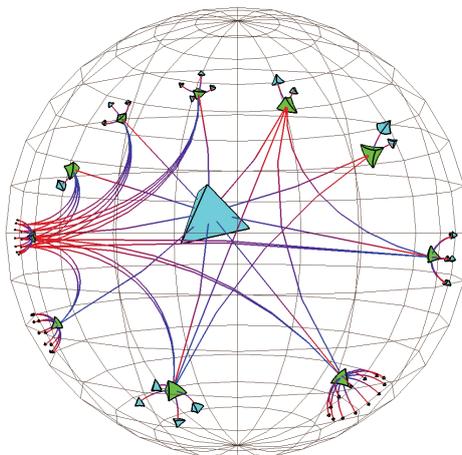


Figure 34 - représentation d'un réseau de pages Internet [Munzner & Burchard, 1995].

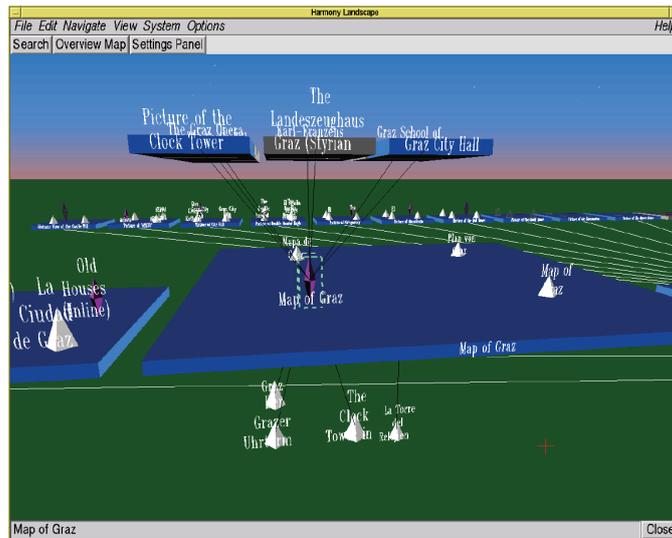


Figure 35 - “Harmony Information Landscape” [Andrews, 1995] : représentation de graphes en 3D.

Toutes ces illustrations mettent en évidence le problème principal des représentations nœud-lien : le problème d’occlusion. Il est dû à l’enchevêtrement des liens, dès que la taille du graphe ou la densité des liens augmente un peu. Il devient alors très difficile pour l’utilisateur d’explorer le graphe visuellement et d’interagir avec ses éléments constitutifs.

Face à de grands graphes, les techniques de représentations matricielles sont des solutions alternatives très intéressantes pour contourner ce problème d’occlusion.

Représentations matricielles

Les représentations matricielles de graphes reposent sur leur matrice d’adjacence. Il s’agit de matrices de booléens (initialisés à faux) où les sommets du graphe sont disposés en lignes et en colonnes. Pour chaque tuple du graphe, l’intersection de la ligne et de la colonne correspondante contient la valeur « vrai ».

La représentation d’une matrice est faite sous forme de grille comme dans l’illustration ci-dessous avec un graphe possédant une seule relation binaire :

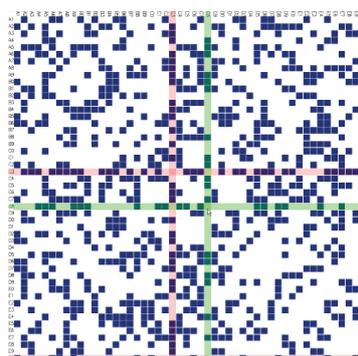


Figure 36 - Les matrices : représentation de graphes [Ghoniem *et al.*, 2004].

Pour augmenter la richesse de la représentation, il est possible de remplacer les booléens par des valeurs numériques et ainsi, il devient possible de représenter des relations binaires valuées [Ghoniem *et al.*, 2004].

Un des principaux avantages de ce type de représentation est de permettre d'éviter les problèmes d'occlusion [Ghoniem *et al.*, 2004]. De plus, comme autre avantage important, les représentations matricielles permettent d'identifier des modèles (structures sous-jacentes) dans l'espace informationnel sous réserve d'appliquer un certain nombre de permutations entre les lignes et les colonnes [Bertin, 1999 (1re éd. 1969)]. Cette propriété est particulièrement intéressante pour le datamining.

En revanche, l'inconvénient majeur de ce type de représentation est sa pauvreté. En effet, il est très délicat de représenter plusieurs relations non binaires avec une seule matrice. De plus, tout comme les représentations en coordonnées parallèles, les représentations matricielles ne permettent pas facilement d'appréhender chaque entité indépendamment des autres.

Ces inconvénients expliquent certainement pourquoi les représentations matricielles demeurent peu utilisées.

iii. Représentations arborescentes

Une arborescence est un graphe connexe acyclique (sa forme évoque la ramification des branches d'un arbre). Une arborescence est donc un graphe particulier dont chaque élément possède au plus un père. Par conséquent, toutes les techniques pour représenter des graphes sont utilisables avec des arborescences (voir la section précédente).

Représentations en liste indentée

La représentation la plus courante des arborescences est une représentation sous forme de liste indentée. Les éléments de la hiérarchie sont représentés (généralement) par une icône avec le nom de l'élément placé à droite. Chaque élément est situé sous l'élément de niveau hiérarchique immédiatement supérieur avec une indentation plus importante (voir illustration ci-après). Chaque arbre peut être plié ou déplié pour permettre de masquer ou d'afficher ses sous-éléments. Sa simplicité et son usage courant rendent ce type de représentation intuitive et facile à utiliser.

Le système WebTOC est une extension de cette technique [Nation, 1998]. Il enrichit la représentation des éléments de la hiérarchie par un ensemble de structures visuelles permettant ainsi d'augmenter le nombre d'informations de la carte.

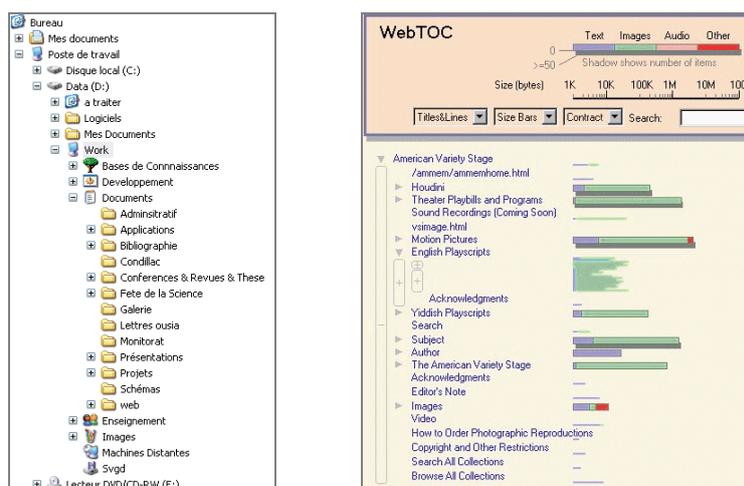


Figure 37 - Liste indentée et le système WebTOC [Nation, 1998] : représentation d'une hiérarchie.

Les représentations sous forme de listes indentées présentent cependant un inconvénient majeur pour représenter des arborescences de grande taille : elles ne parviennent pas à

donner une vision d'ensemble de l'espace informationnel. Le nombre d'éléments affichés de la hiérarchie est limité par l'espace de la carte.

Représentations « nœud-lien » d'arbres

Il existe de nombreux travaux décrivant des paradigmes dédiés à la représentation d'arbres [van Wijk *et al.*, 2003]. Cependant les arbres étant des graphes particuliers, il est possible d'utiliser les techniques de représentation nœud-lien décrites précédemment.

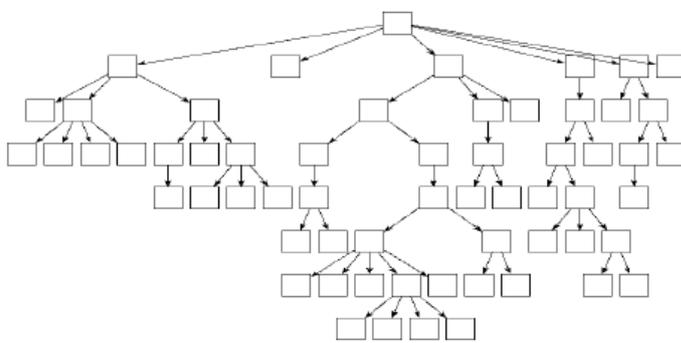
Comme pour les graphes en général, la représentation d'arbres passe par la représentation de ses entités (les nœuds) et leur répartition dans l'espace de la carte. Cette dernière exploite la variable visuelle de position. Le résultat du positionnement dépend alors de la géométrie utilisée. En cartographie d'arborescences deux géométries sont utilisées : la géométrie euclidienne (la plus courante) et la géométrie hyperbolique.

Les paradigmes de représentation d'arbre sont présentés dans les sections suivantes selon la géométrie qu'ils utilisent.

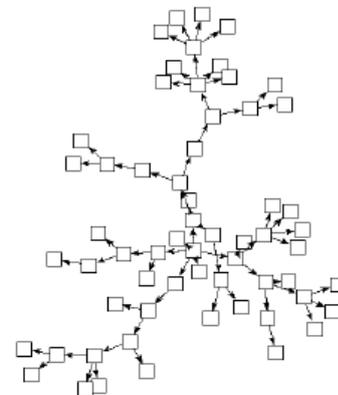
Géométrie euclidienne

En géométrie euclidienne, la principale difficulté pour représenter les hiérarchies est de déterminer le meilleur placement des entités de l'arborescence. Pour cela, de nombreux algorithmes existent [Herman *et al.*, 2000]. La plupart de ces algorithmes ont pour objectif de mettre en valeur la structuration hiérarchique de l'espace informationnel et/ou d'optimiser l'espace de la carte.

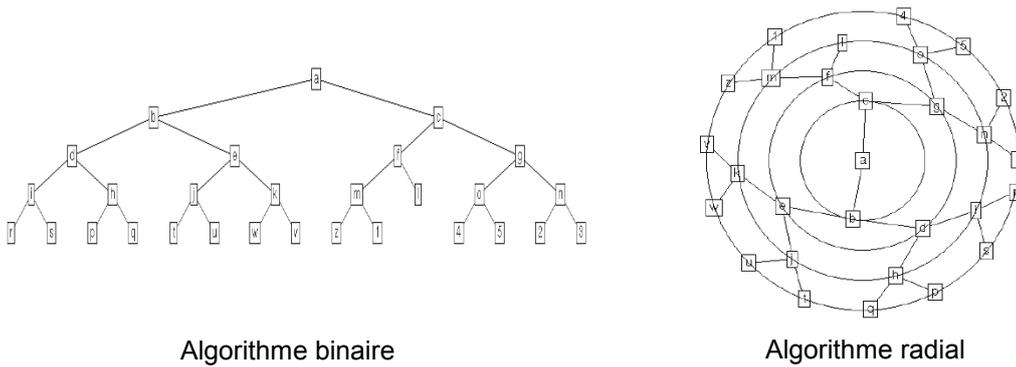
Les schémas ci-après correspondent aux résultats obtenus avec quelques algorithmes de dessin d'arbres :



Algorithme de Walker



Algorithme de Kamada



Algorithme binaire

Algorithme radial

Figure 38 - Quelques algorithmes de dessin d'arbres
[Hascoët & Beaudouin-Lafon, 2001].

Toujours dans le cadre de la géométrie euclidienne, des techniques cartographient les entités de la hiérarchie dans un espace à trois dimensions plutôt que deux. Le plus connu de ces systèmes est celui des arbres de cônes [Robertson *et al.*, 1991].

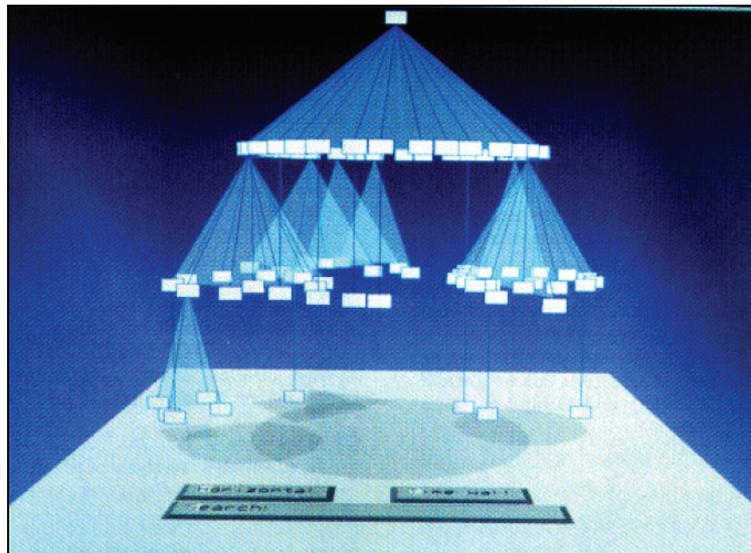


Figure 39 - Les arbres de cônes : représentation d'une hiérarchie dans un espace à trois dimensions [Robertson *et al.*, 1991].

L'utilisation d'une telle représentation implique certains désagréments et le plus dommageable est celui de faire évoluer l'utilisateur dans un espace virtuel en trois dimensions. Il est alors confronté aux problèmes d'occlusion, de navigation et d'orientation.

Arbres hyperboliques

Face à de grandes hiérarchies, l'espace d'une carte est souvent trop restreint et il faut alors choisir entre niveau de détails et vision globale. Pour répondre en partie à cette problématique, certains travaux proposent de répartir les entités radialement dans un espace possédant une géométrie hyperbolique [Lamping *et al.*, 1995].

À l'origine, les auteurs de ces travaux ont eu cette idée en observant le tableau de Maurits Cornelius Escher suivant :

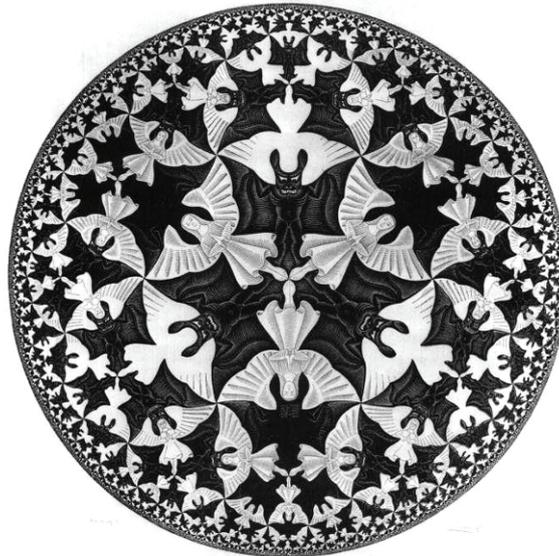


Figure 40 - « Circle Limit IV » de Maurits Cornelius Escher¹.

La géométrie hyperbolique est fondée sur la négation du cinquième postulat d'Euclide : par un point, il passe plusieurs droites parallèles non sécantes. Le modèle de Poincaré est une application de cette géométrie à un disque. Ainsi, on peut définir la géométrie hyperbolique de la manière suivante : le plan est toujours l'intérieur strict d'un disque dont la frontière est appelée horizon. Les droites sont les arcs de cercles orthogonaux à la frontière du disque².

Les propriétés de la géométrie hyperbolique ont pour conséquence que la distance visuelle entre deux points tend vers zéro lorsqu'ils se rapprochent de la frontière du disque. Ainsi, la carte donne l'impression que les structures visuelles sont magnifiées au centre. Inversement, plus elles sont éloignées du centre plus elles sont réduites.

Dans le cas d'une arborescence, les nœuds sont répartis radialement et l'utilisateur a l'impression que la taille des nœuds et la distance entre chaque nœud sont inversement proportionnelles à leur distance au centre du disque. Ainsi, les nœuds sont toujours visibles sinon accessibles, et il suffit à l'utilisateur de glisser au centre ceux qu'il souhaite voir plus en détails.

Voici un exemple avec l'image ci-dessous représentant un arbre avec une géométrie hyperbolique :

¹ M.C. Escher's "Circle Limit IV" (c) 2006 The M.C. Escher Company – the Netherlands. All rights reserved. Used by permission. www.mcescher.com

² <http://www-cabri.imag.fr/abracadabri/abrajava/GNECJ/HDrt01.html>

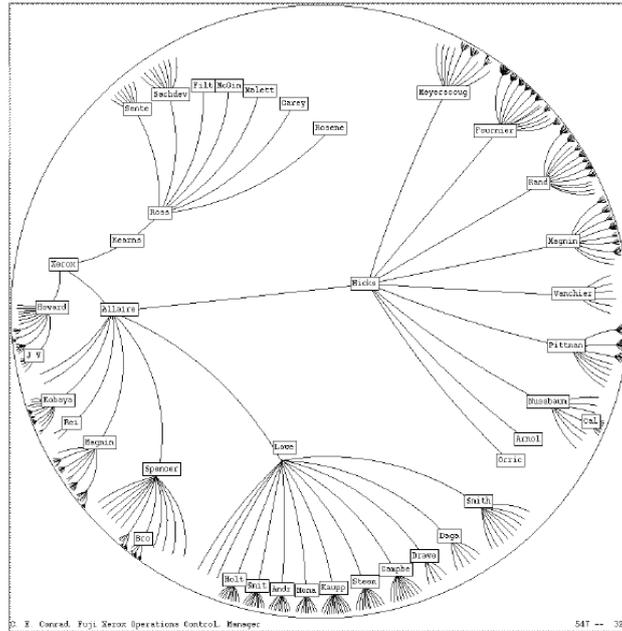


Figure 41 - Navigateur hiérarchique hyperbolique : représentation d'une hiérarchie à l'aide d'une géométrie hyperbolique [Lamping *et al.*, 1995].

Les arbres hyperboliques sont recommandés pour la recherche d'informations dans une arborescence [Pirolli *et al.*, 2003]. En effet, ils ont la capacité d'aider les utilisateurs à « deviner et anticiper » les sous-éléments de chaque élément.

Si de prime abord la forte interactivité des arbres hyperboliques séduit, elle souffre de plusieurs défauts qui peuvent en limiter sa réelle utilisation. Dû aux effets de la déformation, les étiquettes associées aux nœuds ne sont pas alignées et parfois se superposent. Mais c'est principalement son utilisation qui pose problème. En effet, lors de la manipulation de la structure, les éléments à la frontière de l'espace de visualisation se retrouvent projetés de façon « imprévisible ». Ces effets ont tendance à perturber l'utilisateur qui cherche en permanence à rétablir la situation engendrant un effort cognitif plus important et une prise en main assez délicate. Ces effets de projection sont dus à la géométrie utilisée. En effet, les éléments sont représentés dans un plan hyperbolique qui n'est pas commun à nos sens. C'est pourquoi, le résultat des transformations appliquées au plan n'est pas prévisible « naturellement ».

Par la suite des travaux ont étendu les techniques de cartographie avec une géométrie pour passer d'un espace de carte de deux à trois dimensions comme dans l'exemple suivant :

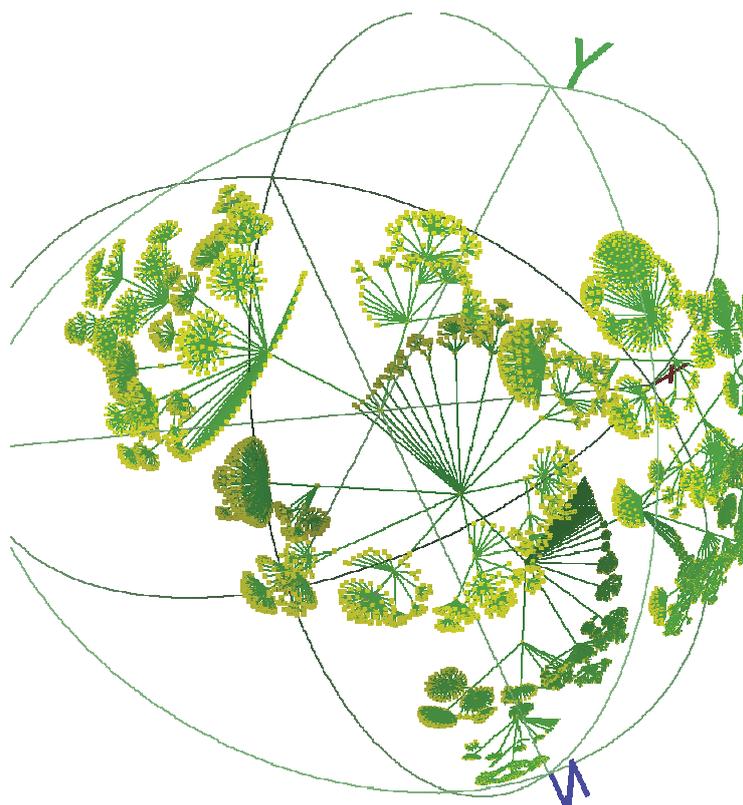


Figure 42 - Navigateur hiérarchique hyperbolique 3D : représentation d'une hiérarchie à l'aide d'une géométrie hyperbolique [Hughes *et al.*, 2004].

Ces deux sections présentent donc les principales techniques de représentation d'arborescences de type nœud-lien. Il nous reste alors à présenter les techniques par pavage.

Représentations par pavage

Les techniques de représentation d'arborescences par pavage (ou parfois aussi nommée surfacique) sont à l'origine des techniques développées pour permettre de visualiser de grands arbres tout en évitant les problèmes d'occlusion (comme c'est parfois le cas avec les représentations de type nœud-lien). Le principe de ces techniques consiste alors à occuper au maximum l'espace de la carte par des structures visuelles surfaciques superposées.

La plus ancienne et la plus connue de ces techniques est celle des Tree-Maps.

Les Tree-Maps et ses variantes

Les Tree-Maps sont issus des premiers travaux qui proposaient une approche par pavage [Baker & Eick, 1995; Johnson & Shneiderman, 1991; Shneiderman, 1992]. Leur principe consiste à découper en surface rectangulaire l'espace de la carte proportionnellement à chaque sous-arbre. Les rectangles visibles correspondent aux feuilles de l'arbre. Chaque nœud de l'arbre est représenté par un rectangle englobant l'ensemble de ses sous-éléments. Ainsi chaque élément subsume graphiquement ses sous-éléments.

Voici un schéma d'un même arbre représenté en technique nœud-lien (à gauche) et en technique Tree-Maps (à droite) :

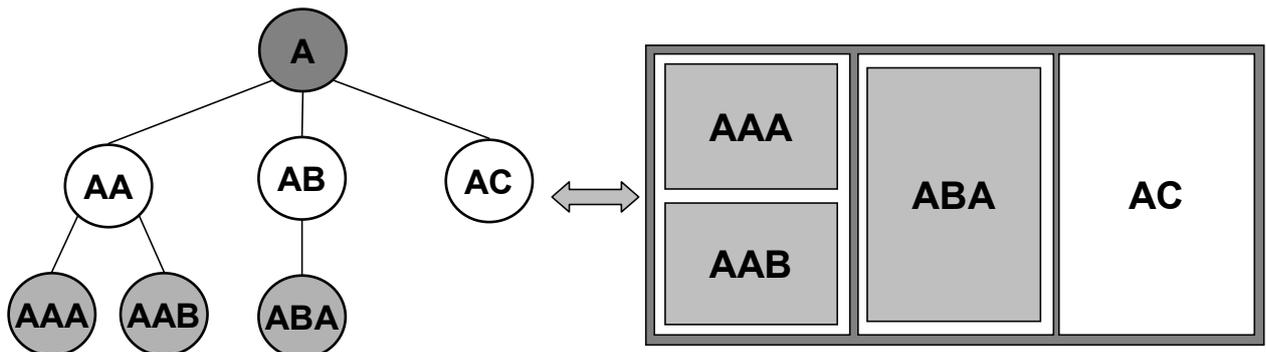


Figure 43 - Comparaison en représentation nœud-lien et Tree-Maps.

Avec les représentations par pavage, la structure n'est pas représentée explicitement et l'attention de l'utilisateur est focalisée sur les feuilles de l'arbre (AAA, AAB et AC dans notre exemple).

Cette caractéristique est plus flagrante avec une saisie d'écran de la première réalisation des Tree-Maps [Johnson & Shneiderman, 1991] :

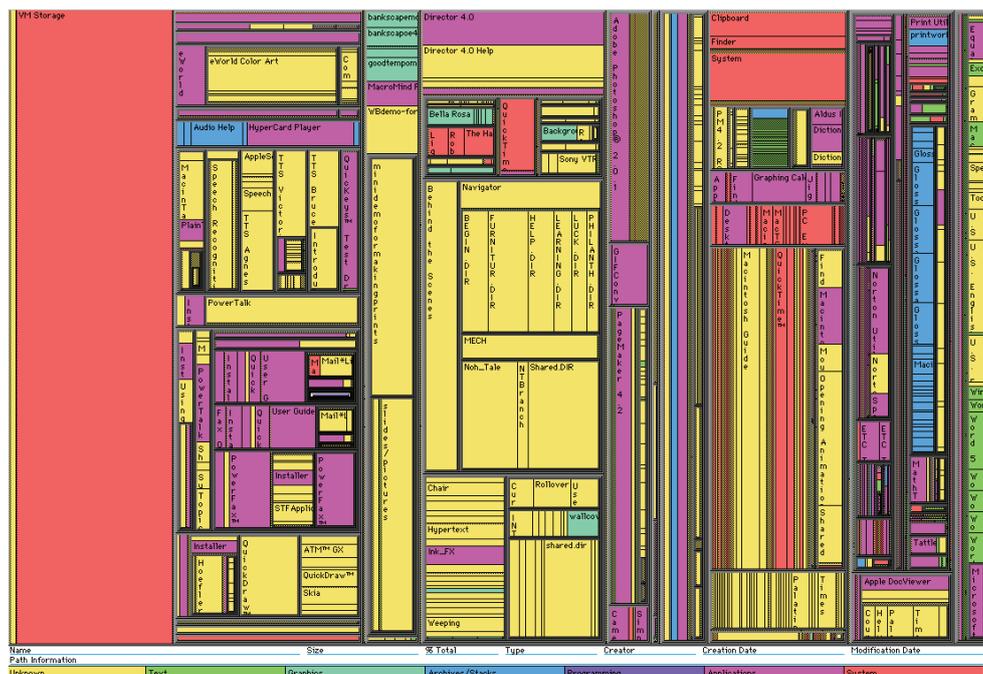


Figure 44 - Tree-Maps : représentation de hiérarchies avec une technique par pavage [Johnson & Shneiderman, 1991]

La non-représentation explicite de la structure est l'inconvénient majeur de cette technique car l'utilisateur ne peut pas la percevoir simplement. Le parcours visuel de la structure nécessite une lecture approfondie de la carte entraînant une surcharge cognitive.

Pour contourner ce problème, certains travaux ont étendu le principe des Tree-Maps avec pour objectif de mettre d'avantage en évidence la structure de l'arborescence. La première de

ces évolutions est celle des « Cushion¹ Treemaps » qui rajoute une impression de relief sur les rectangles.

Voici une saisie d'écran de ce système :



Figure 45 - Cushion Treemaps : évolution des Tree-Maps avec une impression de relief [van Wijk & van de Wetering, 1999].

L'appartenance d'un « rectangle » à un « rectangle » de niveau hiérarchique supérieur est plus facile à détecter.

La technique des « Beamtrees » est une autre évolution des Tree-Maps. Il s'agit cette fois de représenter les nœuds de l'arbre par des cylindres et de placer chaque sous-nœud perpendiculairement par rapport à son père. L'ensemble constitue une carte en trois dimensions. L'objectif est encore ici de faciliter la perception de la structure.

Voici une saisie d'écran de ce système :

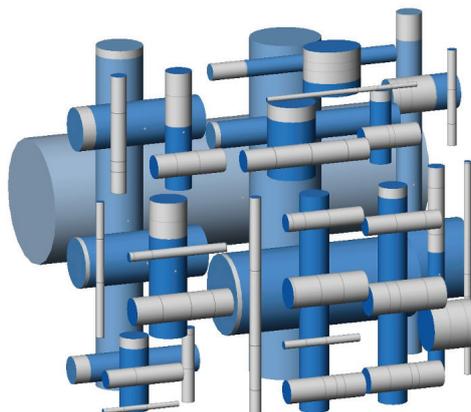


Figure 46 - Beamtrees : évolution des Tree-Maps en 3D [van Ham & van Wijk, 2003].

¹ Cushion signifie coussin en anglais.

Vue de face, la carte est une projection en deux dimensions et l'on obtient une carte similaire au « Cushion Treemaps » mais avec les rectangles (projection des cylindres de face) placés perpendiculairement les uns par rapport aux autres.

Le métamoteur de recherche sur Internet Grokker¹ propose une cartographie qui reprend le même principe que les Tree-Maps mais, avec des surfaces circulaires plutôt que rectangulaires. Ce système est utilisé pour représenter une hiérarchie de « clusters » de sites Internet lors d'une recherche.

Voici une saisie d'écran de ce système lors d'une recherche avec la requête « knowledge visualization » :

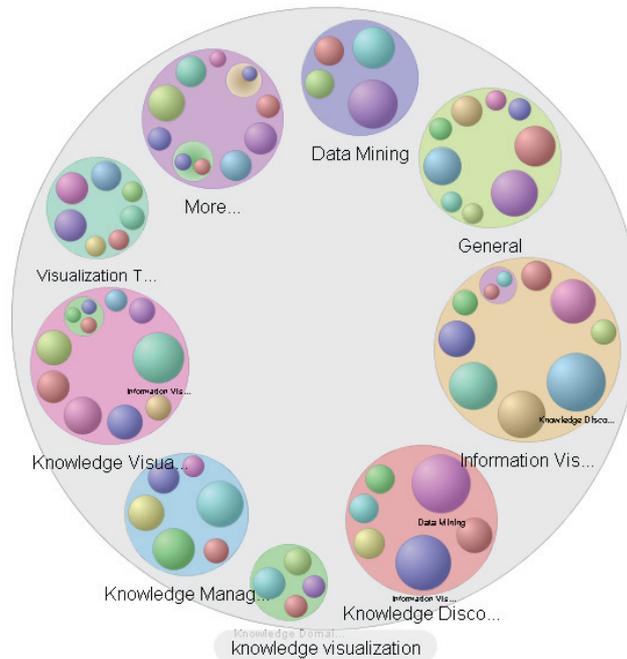


Figure 47 - Grokker : technique de représentation par pavage de cercles.

Les disques d'informations

Les représentations en demi-disques reprennent le principe des techniques par pavage [Andrews & Heidegger, 1998]. Chaque niveau de l'arbre est représenté par un anneau. Les nœuds de l'arbre sont ensuite répartis sur ces anneaux en fonction de leur niveau et en fonction de la position de leur père. Ils sont représentés par des portions d'anneaux.

Voici une saisie d'écran de ce système :

¹ <http://www.grokker.com>

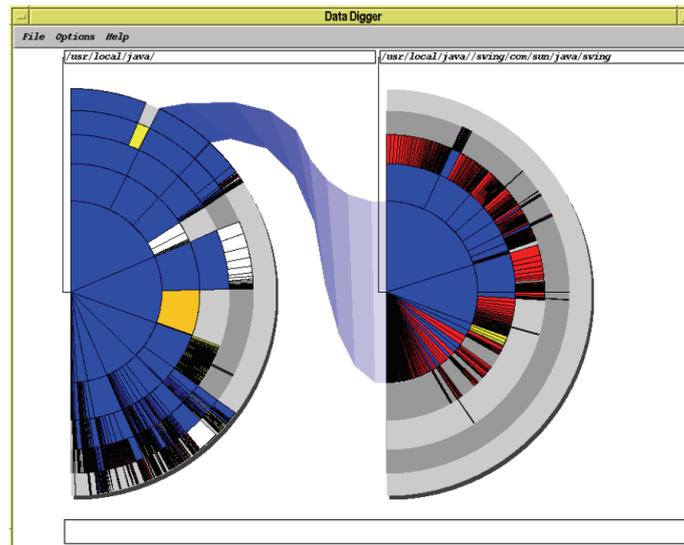


Figure 48 - Les disques d'informations : représentation par pavage d'une hiérarchie [Andrews & Heidegger, 1998].

L'avantage de cette technique par pavage est qu'elle permet de mieux percevoir la structure arborescente des éléments que les Treemaps. Cependant, la structure n'est pas représentée explicitement et son parcours visuel n'est pas toujours aisé. Néanmoins, cette représentation donne facilement une interprétation ensembliste de la structure.

Le système « Sunburst » offre une évolution du système précédent en répartissant les nœuds sur 360° et en ne colorant pas les anneaux correspondant aux niveaux intermédiaires [Stasko & Zhang, 2000].

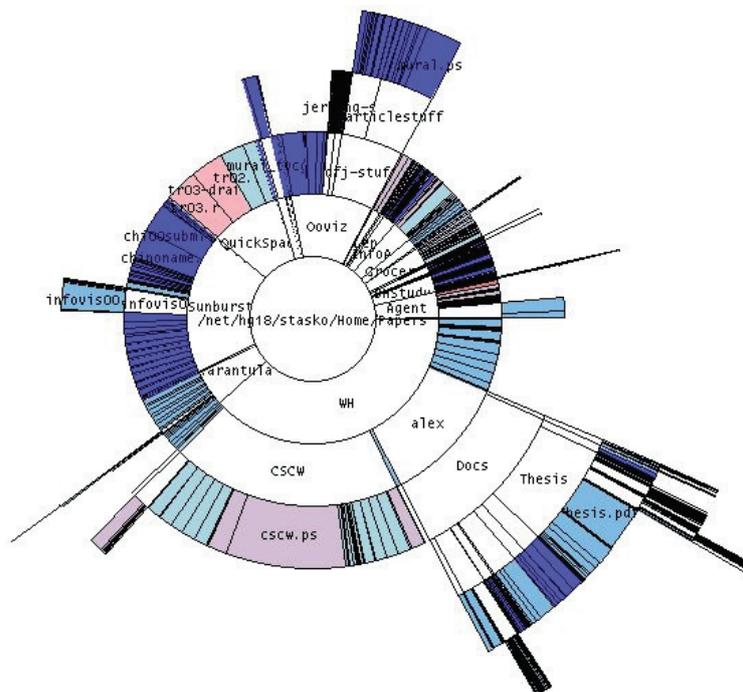


Figure 49 - Sunburst : représentation par pavage d'une hiérarchie [Stasko & Zhang, 2000].

Métaphore virtuelle

D'autres travaux présentent des solutions pour représenter une hiérarchie avec une métaphore virtuelle. Voici deux exemples de ces systèmes : le cube d'informations et les arbres botaniques.

Le cube d'information consiste à représenter chaque élément de la hiérarchie par un cube imbriqué dans celui qui correspond à son père [Rekimoto & Green, 1993]. La perception des éléments les plus imbriqués est garantie par une gestion de la transparence.



Figure 50 - Les cubes d'informations : représentation de hiérarchie en 3D par imbrication de cubes [Rekimoto & Green, 1993].

Une autre variante consiste à représenter la hiérarchie comme un arbre « botanique » en trois dimensions [Kleiberg *et al.*, 2001]. Les feuilles de la hiérarchie sont représentées par des fruits ce qui permet d'enrichir la représentation avec l'exploitation de nombreuses variables visuelles.

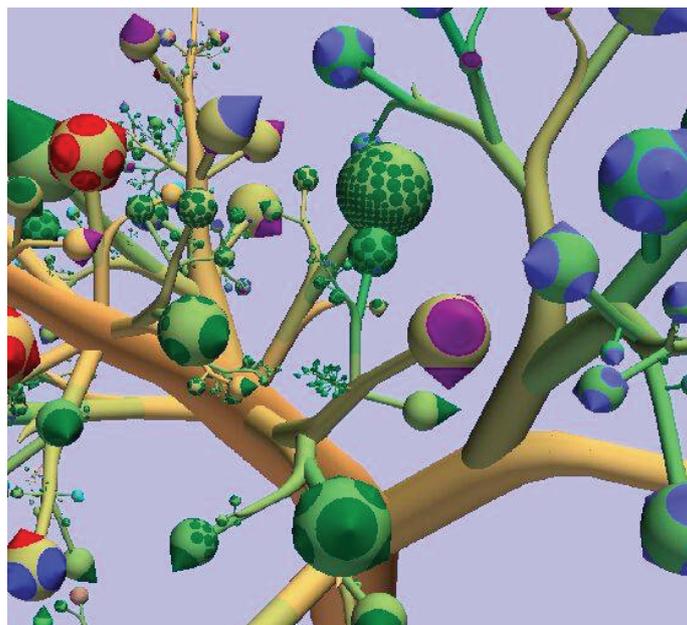


Figure 51 - Visualisation "botanique" : représentation de hiérarchies avec une métaphore virtuelle en 3D [Kleiberg et al., 2001].

Dans ce système, les nœuds intermédiaires ne sont pas représentés explicitement. Cette représentation met donc en valeur la structure hiérarchique par opposition au système précédent. Néanmoins, comme toutes les représentations en 3D, l'utilisateur est confronté aux problèmes d'occlusion.

b. Les paradigmes de visualisation

Les paradigmes de visualisation permettent de construire un nouvel espace correspondant à la visualisation de l'espace informationnel *représenté* (ou représentation). Nous nommons ce nouvel espace un « espace informationnel *visualisé* » ou plus simplement, une « visualisation ».

La construction d'une visualisation à partir d'une représentation correspond à l'action de « visualiser ».

Remarque : le sens de « visualiser » ne correspond pas à l'action de voir – percevoir par le sens de la vue – mais bien à celui de « rendre sensible à la vue ce qui n'est pas naturellement visible »¹.

L'objectif de l'action de visualiser est donc de rendre accessible au sens de la vue les structures visuelles de la représentation. Par exemple, pour visualiser un texte il est possible de l'imprimer ou bien encore, pour visualiser une carte, il est possible de l'afficher sur un écran d'ordinateur.

Avec ces deux exemples, « visualiser » semble n'être qu'une opération assimilable à une « projection » d'une représentation sur un support. Mais ce point de vue est très réducteur car, il existe de nombreux paradigmes qui permettent de passer d'une représentation à une

¹ D'après le Trésor de la Langue Française.

multitude de visualisations très différentes. Citons par exemple, les techniques de mise en perspective du plan de la carte en vue de présenter un plus grand nombre d'informations.

Il existe deux grandes classes de techniques de visualisations :

- > Les techniques de visualisations uniformes : ces techniques permettent d'avoir une projection sans déformation de l'espace de la carte (avec éventuellement des transformations affines).
- > Les techniques de visualisations non uniformes : ces techniques permettent de projeter l'espace de la carte avec une ou plusieurs déformations permettant de visualiser un plus grand nombre de structures visuelles.

Ces deux classes de techniques de visualisations sont décrites dans les deux sections suivantes.

i. Les visualisations uniformes : contrôle du point de vue

Les visualisations uniformes sont les visualisations les plus courantes. Elles permettent de visualiser sans déformation (donc uniforme) tout ou partie d'un espace informationnel représenté.

Dans l'espace informationnel représenté, les structures visuelles possèdent une position (définie par des coordonnées sur les axes de la carte). Toutes les structures définissent ainsi l'espace de la représentation qui les contient toutes.

L'espace de la carte, correspondant à la zone d'affichage (comme la taille de l'écran ou de la feuille), est limité et différent de celui de la représentation (qui englobe toutes les structures visuelles). La visualisation consiste alors à visualiser (uniformément) l'espace représenté en l'adaptant à l'espace de la carte. Dans le cadre des visualisations uniformes, les techniques utilisées sont des transformations appliquées à l'espace de la représentation (défini par les structures visuelles). Parmi ces transformations, les plus couramment utilisées sont des transformations affines comme l'application d'un facteur de zoom, une translation ou une rotation ou bien encore, des transformations comme la sélection d'un espace plus restreint. Toutes ces transformations affines permettent de définir un point de vue sur l'espace de représentation.

Par exemple avec l'illustration suivante, deux transformations sont appliquées à l'espace de représentation : une translation (déterminée par la sélection) puis une homothétie (centrée à l'origine de la figure).

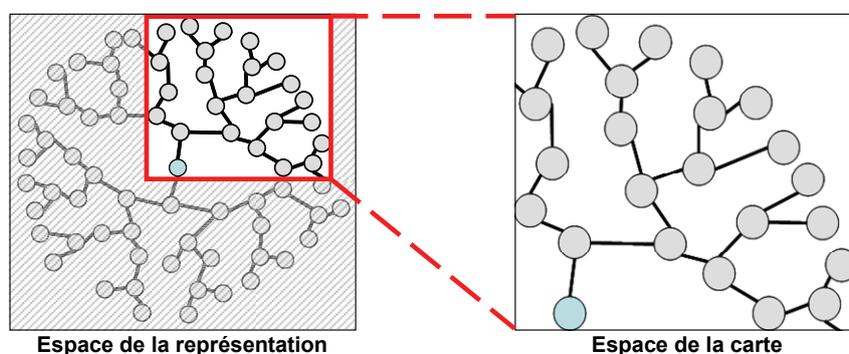


Figure 52 - Exemple de contrôle de point de vue avec deux transformations affines.

Lors de l'utilisation d'une carte, il est généralement nécessaire d'avoir une vision globale de l'ensemble de l'espace informationnel mais aussi de voir les éléments de la carte en détails. Pour cela, l'utilisateur est contraint de changer en permanence de point de vue (par exemple en changeant de facteur de zoom). Il en résulte une surcharge cognitive et bien souvent, les utilisateurs finissent par être déstabilisés et par perdre le chemin qu'ils ont parcouru dans la carte. Pour mieux comprendre le principe des navigations basées sur des successions de changements de point de vue, une modélisation en est proposée dans [Furnas & Bederson, 1995]. De plus, un système de métriques est proposé dans [van Wijk & Nuij, 2003] pour calculer une animation optimale entre deux changements de point de vue.

Pour éviter l'inconvénient de la surcharge cognitive, il est possible d'utiliser des vues multiples d'une même représentation [Baldonado *et al.*, 2000; North & Schneiderman, 1997]. Elles sont alors liées et coordonnées. La mise en œuvre de vues multiples répond à certaines règles comme la complémentarité (les vues ont pour but de mettre en évidence les similarités ou les différences) ou encore la décomposition (les vues ont pour objectif de subdiviser les éléments à représenter pour mieux les appréhender).

De nombreux systèmes de visualisations exploitent cette technique en utilisant deux vues : la première donne un point de vue global mais peu détaillé de l'ensemble des éléments de la carte et la seconde, donne une vision détaillée des éléments directement dans la zone de focalisation de l'utilisateur. Cette technique est nommée « overview + detail » [Card *et al.*, 1999c; Plaisant *et al.*, 1995].

L'image ci-après est une saisie d'écran issue des travaux nommés « informations murales » [Jerding & Stasko, 1995] portant sur l'étude des taches solaires par mois depuis le XVIIIe siècle. La vue supérieure correspond à la vue détaillée de la zone rectangulaire sélectionnée (en rouge) dans la vue inférieure.

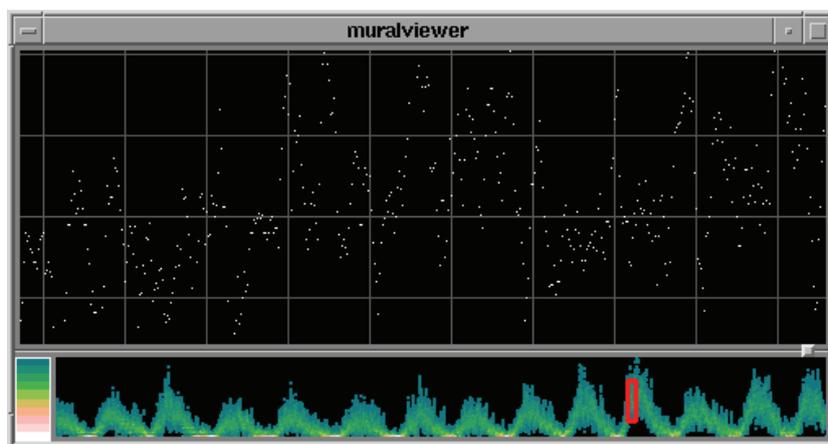


Figure 53 - Exemple de vues multiples : principe "overview + detail" [Jerding & Stasko, 1995].

Bien que le paradigme des vues multiples soit efficace, il ne faut pas négliger les effets négatifs ; le principal étant la création d'une discontinuité spatiale avec plusieurs points de focalisation, ce qui entraîne une surcharge cognitive pour les utilisateurs.

ii. Les visualisations non uniformes : focalisation et contexte

L'espace de la carte étant par nature restreint et le nombre d'informations à cartographier toujours plus important, il est très difficile d'afficher toutes les informations avec un maximum de détails. Les opérations de contrôle de point de vue (comme les changements de zoom)

permettent de résoudre en partie ce problème. Cependant, lors de changements de point de vue, l'utilisateur est vite déstabilisé.

Cette problématique est à l'origine des techniques de visualisations non uniformes. Elles ont pour objectif de remédier au problème du manque d'espace par rapport à la quantité d'information à cartographier.

Le principe commun de toutes ces techniques est d'afficher les éléments de la carte avec un niveau de détails variable en fonction de l'intérêt que leur porte l'utilisateur de la carte. Elles permettent donc de tenir compte d'un certain nombre de points d'intérêt de l'utilisateur nommés focalisations ou « focus ».

Dans les zones de focalisation, les structures visuelles sont visualisées de manière optimale et pour le reste de la carte, elles sont visualisées sous forme « dégradée » (plus « abstraite »), mais néanmoins informante [Furnas, 1981; Leung & Apperley, 1994; Sarkar & Brown, 1992]. Le résultat de ce type de visualisation est une vue dite non uniforme.

Pour illustrer ce principe prenons l'image ci-après extraite des albums de bande dessinée « Astérix et Obélix » :

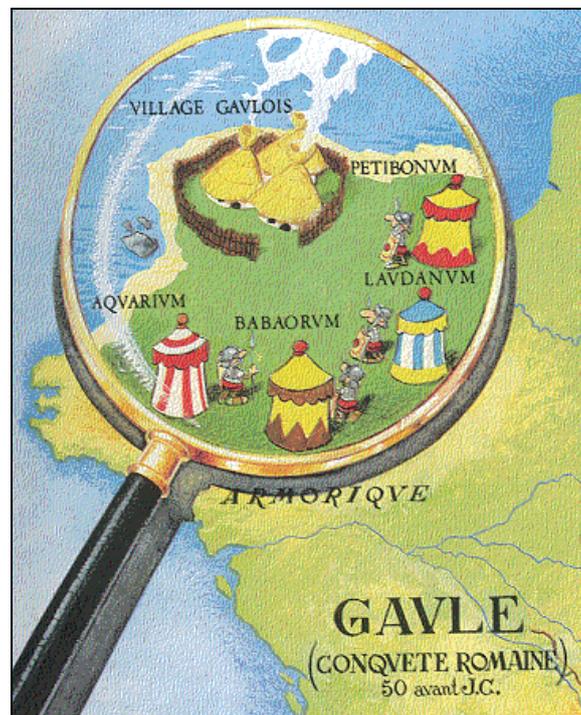


Figure 54 - Astérix et Obélix : exemple de visualisation non uniforme.

Dans cette illustration, la visualisation n'est pas uniforme. Le célèbre village gaulois est magnifié (taille plus importante et niveau de détails plus élevé) alors que le reste de la carte est diminué (vision globale et pauvre en détails).

Une vue non uniforme est obtenue en appliquant une fonction, nommée « fonction de transformation », à l'espace *représenté* [Leung & Apperley, 1994]. L'illustration ci-dessous présente l'application d'une fonction de transformation à une structure visuelle de l'espace *représenté* :

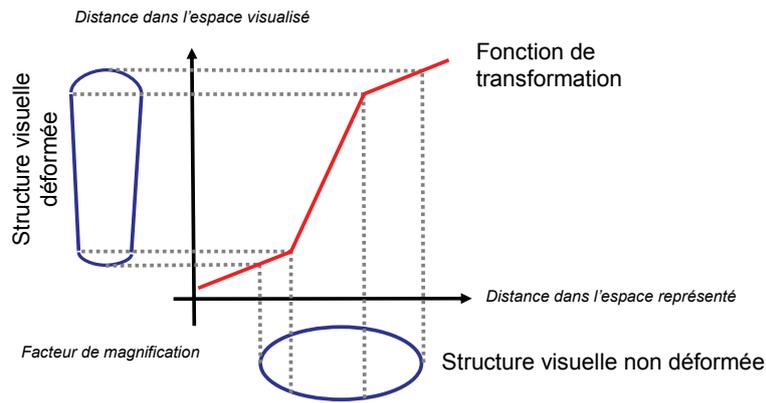


Figure 55 - Vues non uniformes : la fonction de transformation permet de déformer les structures visuelles [Leung & Apperley, 1994].

La dérivée de la fonction de transformation nommée « fonction de magnification » fournit le profil de la transformation (amplification ou réduction) et plus particulièrement, un facteur de magnification (correspondant à la « force » de la déformation).

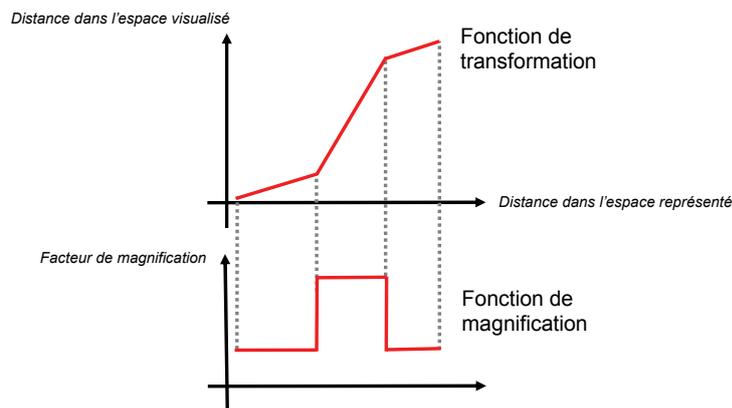


Figure 56 - Vues non uniformes : la fonction et le facteur de magnification [Leung & Apperley, 1994].

Il existe trois grandes classes de techniques pour produire des vues non uniformes : les visualisations bifocales (comme les lentilles documentaires et les murs fuyants), les visualisations polyfocales (comme les tableaux optiques) et les visualisations « fisheye ».

Ces trois classes sont décrites dans les sections suivantes.

Les visualisations bifocales et leurs variantes

Le paradigme des visualisations bifocales est une technique intuitive et aisée à mettre en œuvre [Spence & Apperley, 1999]. Cette technique a une fonction de transformation linéaire unique pour la zone de focalisation et pour le reste de la carte. Cette fonction peut être appliquée à une ou plusieurs dimensions de l'espace *représenté* (sur un ou plusieurs axes de la carte).

Dans les exemples ci-dessous, une fonction de transformation similaire au schéma précédent est appliquée à un quadrillage uniforme :

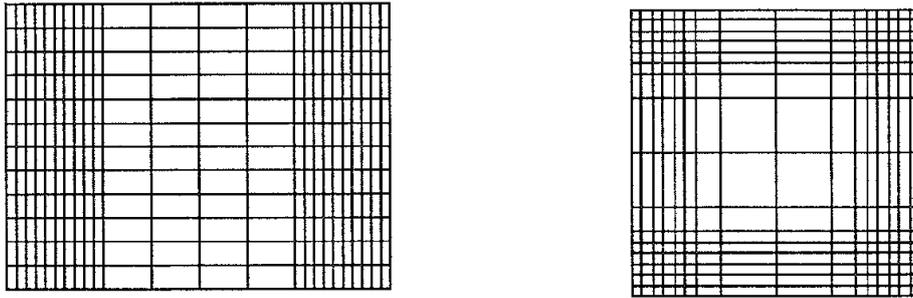


Figure 57 - Application d'une transformation bifocale sur une et deux dimensions de la carte [Leung & Apperley, 1994].

Dans le premier cas, la fonction de transformation est appliquée à une seule dimension (axes des abscisses) par conséquent, la carte est déformée horizontalement. Dans le second cas, la fonction est appliquée aux deux axes de la carte par conséquent, la carte est déformée horizontalement et verticalement.

Les deux sections suivantes décrivent deux systèmes qui exploitent la technique des visualisations bifocales : les lentilles documentaires et les murs fuyants.

Les lentilles documentaires

Le système des « lentilles documentaires » [Robertson & Mackinlay, 1993] est une variante du paradigme de l'affichage bifocal. Dans ce système, la fonction de transformation est appliquée à un pavage de pages de documents.

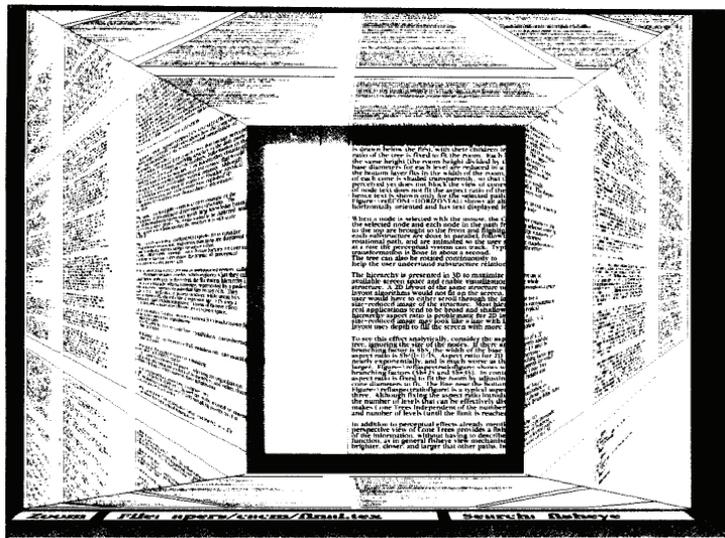


Figure 58 - Lentille documentaire [Robertson & Mackinlay, 1993].

Les murs fuyants

Les murs fuyants (Perspective Wall) constituent une évolution de l'affichage bifocal [Mackinlay *et al.*, 1991]. La partie centrale de la carte n'est pas transformée alors que sur les côtés, elle est transformée pour donner une impression de perspective comme dans la saisie d'écran suivante :

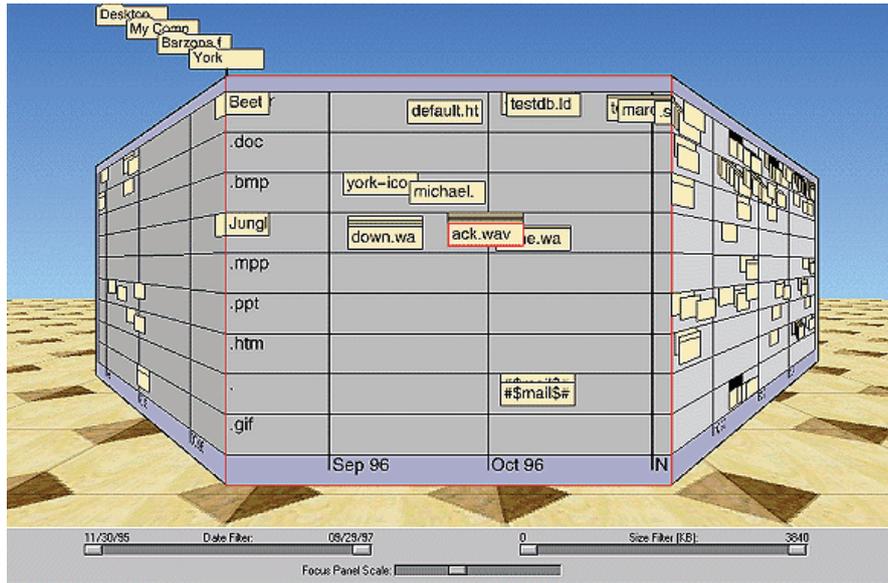


Figure 59 - Mur fuyant : évolution de l'affichage bifocal [Mackinlay *et al.*, 1991].

Les deux panneaux latéraux ont une fonction de magnification « réductrice » proportionnelle à la distance au premier plan de la carte. Ce paradigme est dédié à la visualisation d'espaces vectoriels à une dimension.

Les visualisations polyfocales

Une visualisation poly focale est obtenue en appliquant, à une représentation, une fonction de transformation continue qui possède plusieurs « pic » [Leung & Apperley, 1994; Spence & Apperley, 1999]. La particularité de ce type de transformation est d'obtenir une visualisation avec plusieurs points de focalisation. La plupart des visualisations bifocales sont donc des cas particuliers de visualisations polyfocales.

Voici un exemple de transformation avec plusieurs points de focalisation obtenue à partir d'une grille uniforme :

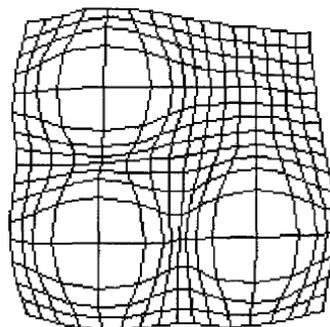


Figure 60 - Application d'une transformation poly focale [Leung & Apperley, 1994].

Le fait d'avoir plusieurs points de focalisation et des techniques d'interactions pour changer les paramètres de la fonction de transformation permet, entre autres, de magnifier et donc de rapprocher des éléments éloignés dans l'espace de la représentation pour pouvoir les comparer.

Le système des tableaux optiques est une mise en œuvre du paradigme de visualisation poly focale appliquée à un tableur comme Excel [Rao & Card, 1994].

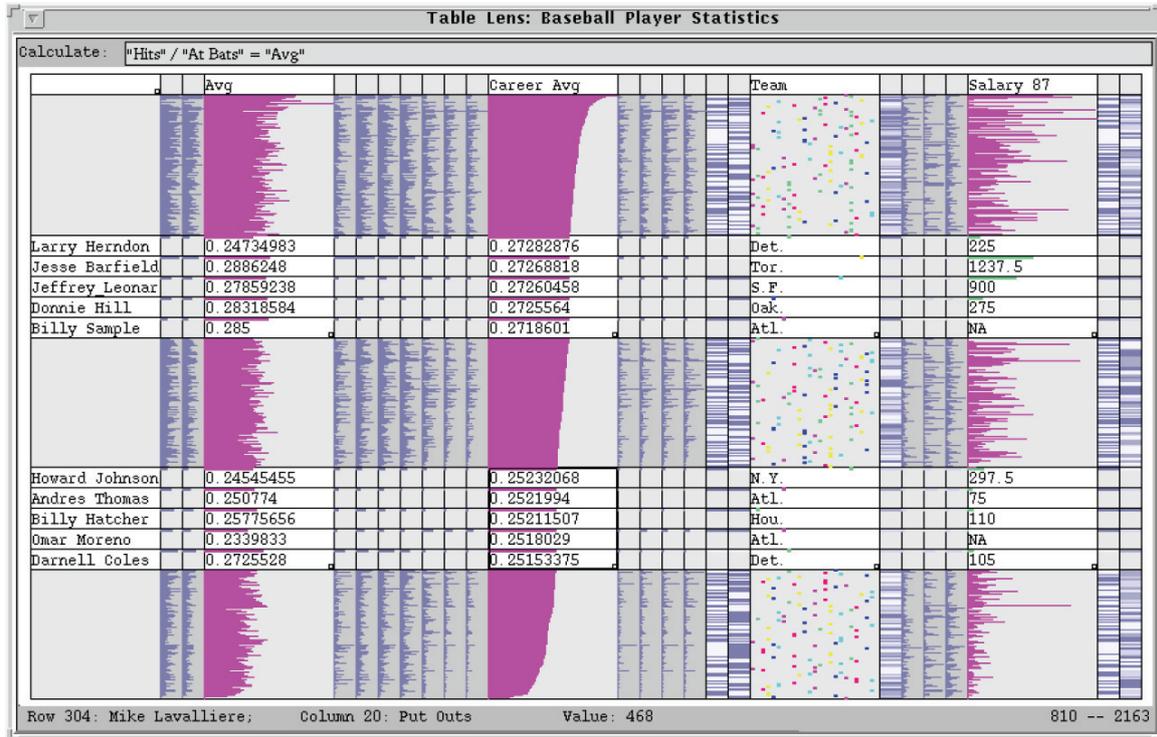


Figure 61 - Tableaux optiques : application de plusieurs points de focalisation sur un tableur [Rao & Card, 1994].

L'utilisateur du système peut rajouter des points de focalisation pour permettre de comparer des plages de cellules très éloignées dans l'espace informationnel représenté.

Les visualisations Fisheye

Les visualisations « Fisheye » (en œil de poisson) permettent de produire des vues en deux dimensions où les structures visuelles de la représentation apparaissent avec des tailles variables.

À l'origine, le fisheye (nommé le « fisheye filtrant ») est une technique pour filtrer l'affichage des informations par rapport à leur degré d'intérêt [Furnas, 1981; Furnas, 1986]. Dans ces travaux, le fisheye est appliqué à des informations textuelles comme un plan de document ou un listing de code source.

Le fisheye permet de calculer une valeur pour chaque élément qui est fonction de l'« importance a priori » de l'élément, diminué de sa distance au centre d'intérêt (le focus).

Voici un exemple de fisheye filtrant appliqué à une arborescence :

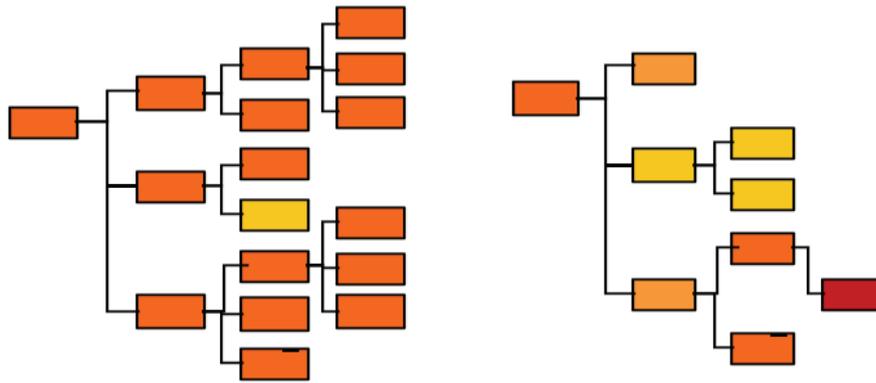


Figure 62 - Fisheye filtrant : application du paradigme du fisheye (dessin de droite) sur un arbre (dessin de gauche) [Hascoët & Beaudouin-Lafon, 2001].

Dans cet exemple, les nœuds qui n'ont pas une valeur suffisante selon le fisheye, ne sont pas affichés.

Par la suite, d'autres travaux ont étendu ce principe pour adapter le paradigme à la graphique [Sarkar & Brown, 1992]. Dans ce cas, aucun élément ne disparaît mais, la taille et la position des structures visuelles sont modifiées. Cette évolution est nommée le « fisheye déformant ».

Tout comme avec les vues bifocales, le fisheye déformant offre un centre de focalisation nommé « centre d'intérêt » et la taille des structures visuelles dans la vue est proportionnelle à la distance à ce point d'intérêt. Certaines structures peuvent alors ne plus être visibles. Le résultat est alors comparable à la vision au travers d'un objectif grand angle d'un appareil photographique : l'utilisateur peut se concentrer sur son centre d'intérêt et utiliser sa vision périphérique (comme celle du poisson ou de la mouche) pour surveiller tout l'espace de la carte. Cette métaphore est d'ailleurs à l'origine du nom de la technique.

Voici un exemple de visualisation fisheye appliquée à un graphe formant une grille uniforme :

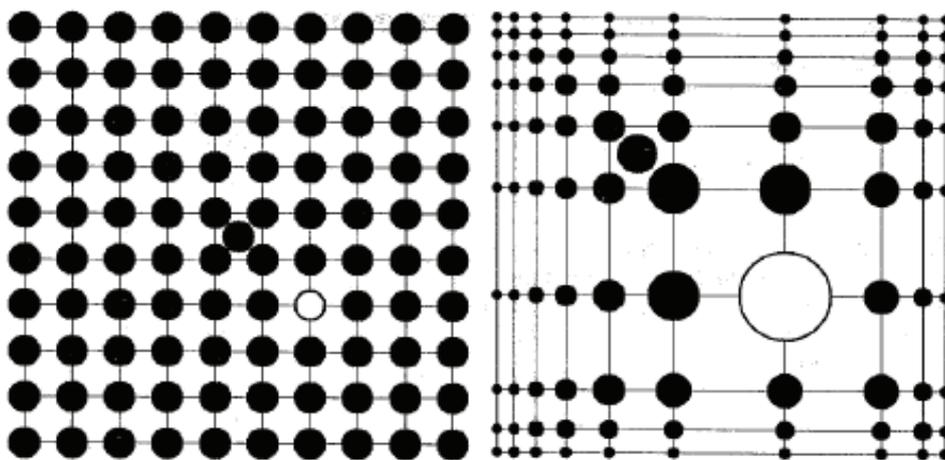


Figure 63 - Fisheye déformant : application du paradigme du fisheye (dessin de droite) sur les coordonnées cartésiennes des éléments d'un graphes (dessin de gauche) [Sarkar & Brown, 1992].

Le nœud blanc correspond au point de focalisation de la vue.

Remarque : le paradigme des arbres hyperboliques et plus particulièrement la géométrie hyperbolique appliquée au disque de Poincaré permet d'obtenir lors de la visualisation un effet fisheye (voir à la page 69).

Les sections précédentes décrivent les différentes techniques permettant de construire une carte à partir d'un espace informationnel. Il nous reste maintenant à décrire les paradigmes d'interaction.

c. Les paradigmes d'interaction

Avec l'outil informatique, les cartes sont devenues interactives. Les utilisateurs ont évolué du statut de « percepteur » à celui d'« acteur ». Parfois même, l'utilisateur détermine les techniques qu'il veut mettre en œuvre et par conséquent, il joue aussi le rôle de cartographe.

On doit cette prodigieuse évolution aux recherches menées dans le domaine des interfaces homme-machine (IHM). Régulièrement, de nouveaux paradigmes d'interaction sont inventés.

Toutes les techniques d'interaction liées à la cartographie ont pour principe fondamental de profiter de la facilité avec laquelle l'homme peut extraire des informations d'un environnement qu'il contrôle directement et activement [Hascoët & Beaudouin-Lafon, 2001]. On parle alors de couplage fort entre l'action et la perception : « Il faut agir pour percevoir et il faut percevoir pour agir ».

Selon la vision en niveaux du processus de cartographie, les interactions sont assimilables à des opérations qui portent sur les différents niveaux de la cartographie [Card *et al.*, 1999b] :

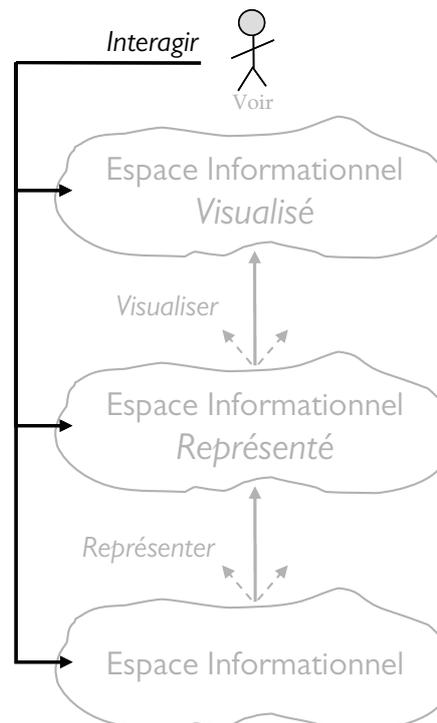


Figure 64 - les interactions dans la cartographie.

Chaque interaction entraîne la modification du niveau sur lequel elle porte [Chi & Riedl, 1998]. Par effet de cascade, toute modification d'un niveau entraîne la mise à jour des niveaux supérieurs.

Par exemple, imaginons que l'utilisateur modifie le nom d'un nœud de la cartographie d'un graphe. Cette modification entraîne la mise à jour de l'espace informationnel. Ensuite, elle provoque la mise à jour de sa représentation (la structure visuelle) puis de sa visualisation.

Les tableaux suivants présentent les principales interactions utilisées en cartographie de données abstraites :

Tableau 3 - Paradigmes d'interaction au niveau de l'espace informationnel.

Technique	Description
Les requêtes dynamiques	<p>Ce type d'interaction permet de soumettre une requête à l'ensemble de l'espace informationnel pour sélectionner uniquement les données qui correspondent au critère de la requête [Ahlberg et al., 1992].</p> <p>Généralement les requêtes sont exprimées par un ensemble de plages de valeurs d'attributs à respecter pour les éléments de l'espace informationnel. Ces plages de valeurs sont spécifiées par l'intermédiaire de composants graphiques comme des ascenseurs, des boutons, ou des cartes cliquables.</p>
Direct walk	<p>Les interactions « direct walk » permettent de changer l'espace informationnel pris en compte pour la cartographie par l'intermédiaire de liens vers d'autres espaces informationnels [Card et al., 1994]. Cette interaction est très courante avec la navigation sur Internet : le passage d'une page à une autre, via des liens hypertextes, est une application de direct walk.</p>
Manipulation directe	<p>Cette manipulation permet de modifier directement les éléments de l'espace informationnel. Par exemple, l'utilisateur sélectionne un élément pour changer sa valeur.</p>

Tableau 4 - Paradigmes d'interaction au niveau de l'espace informationnel *représenté*.

Technique	Description
Détails à la demande	Cette technique permet de se focaliser sur un petit ensemble de structures visuelles afin de révéler davantage de détails sur celles-ci. C'est le cas par exemple avec l'affichage d'une bulle d'information lorsque l'utilisateur survole des éléments avec sa souris. Cette technique permet aussi d'appliquer des transformations équivalentes au zoom sur une carte et donc de représenter plus de données dans une même carte.
Les lentilles magiques	Cette technique permet de sélectionner un ensemble de structures visuelles en fonction de leurs positions dans la carte et de leur appliquer d'autres interactions [Stone <i>et al.</i> , 1994] comme des requêtes dynamiques ou le changement du facteur de zoom. Un exemple courant est celui de la loupe.
Brushing	Cette technique permet de sélectionner un ensemble de structures visuelles dans une vue et de sélectionner automatiquement le même ensemble dans d'autres vues. Cette technique est très courante avec le paradigme des vues multiples (voir ci-dessus page 79).

Tableau 5 - Paradigmes d'interaction au niveau de l'espace informationnel *visualisé*.

Technique	Description
Mouvement de caméra	Cette technique permet de changer la position de l'observateur par rapport aux structures visuelles de la carte. Ce changement de point de vue peut permettre de voir des parties moins visibles de la carte.
Le contrôle de point de vue	Cette technique consiste à permettre à l'utilisateur de changer les différents paramètres du point de vue comme le zoom ou la localisation (voir ci-dessus page 79).

2.3 Synthèse & critiques

Nous voici donc à la fin de notre parcours du domaine de la cartographie de données abstraites. Elle repose sur les travaux de deux communautés : la visualisation d'informations et la visualisation de connaissances.

Notre objectif est d'aider les collaborateurs des organisations à appréhender leur espace informationnel au travers de cartes. D'après l'étude de l'existant, la cartographie de données abstraites permet de répondre en partie à cette problématique. En effet, nous savons que pour construire une carte d'un espace informationnel donné, il est nécessaire de mettre en œuvre des paradigmes de représentation et de visualisation. Nous savons aussi, qu'il faut proposer des paradigmes d'interaction à l'utilisateur pour qu'il puisse interagir avec la carte. Par conséquent, il existe de nombreuses techniques permettant de passer d'un espace informationnel à une quasi-infinité de cartes. Cette « recette » est maintenant bien connue et maîtrisée cependant elle ne répond que partiellement à notre problématique.

Pour rappel, nous avons besoin de proposer une solution de cartographie pour construire des cartes permettant aux utilisateurs de naviguer dans leur espace informationnel selon la sémantique de leur domaine avec une vision adaptée alliant différents niveaux d'échelles.

Pour répondre complètement à ces besoins, il subsiste alors deux problématiques :

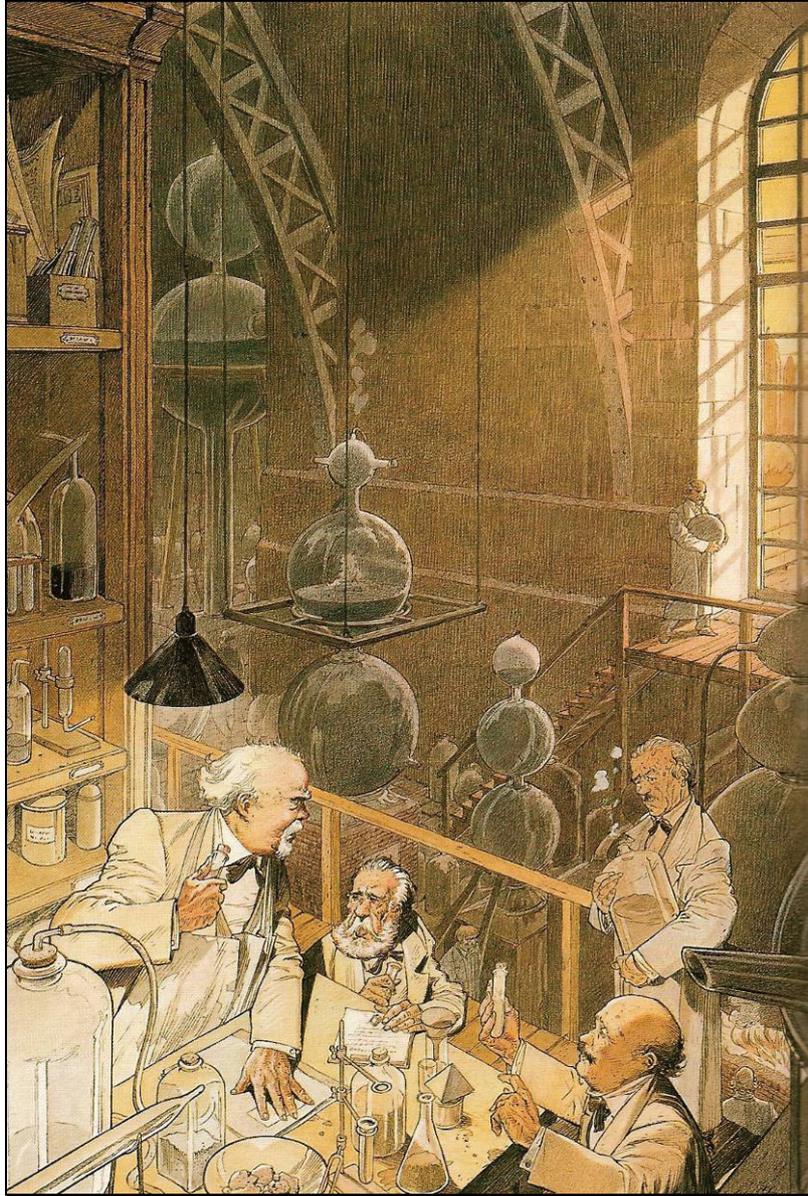
- > Comment proposer une navigation basée sur la sémantique du domaine ?
- > Comment déterminer les paradigmes à mettre en œuvre pour fournir des cartes adaptées à l'utilisateur avec différents niveaux d'échelles ?

Problématique de la navigation sémantique : les cartes doivent permettre aux utilisateurs de naviguer selon la sémantique du domaine. Il est alors nécessaire que cette sémantique soit explicitée dans l'espace informationnel à travers sa modélisation. La difficulté est alors la construction de cette modélisation.

Problématique du choix des paradigmes : la cartographie de données abstraites est assimilable à un processus qui permet de construire une carte à partir d'un ensemble d'informations en employant un certain nombre de techniques. Les différentes approches ne nous indiquent pas comment choisir ces techniques pour construire des cartes alliant différents niveaux d'échelles et une visualisation adaptée à l'utilisateur (adaptée à l'activité et niveau d'expertise ainsi qu'à l'impact des représentations).

Les propositions associées à cette étude ont pour objectif de répondre à cette double problématique.

Partie 3 : **Propositions**



« Devant la masse toujours croissante des données, la cartographie offre des possibilités de communiquer synthétiquement et dégager les informations stratégiques utiles aux décideurs. Seulement, comment passer des données brutes à la carte ? »

Pierre Mongin (préface de [Poidevin, 1999])

Rappel du plan

3.1 La cartographie sémantique	93
3.1.1 Cartographie orientée sémantique	94
3.1.2 Cartographie orientée retours d'expérience	97
3.1.3 Synthèse.....	98
3.2 Notre processus de cartographie sémantique	99
3.2.1 Des données à la carte	99
3.2.2 Des connaissances à la carte	101
3.2.3 Description du processus	103
3.2.4 Synthèse.....	107
3.3 SNDF : notre formalisme de description pour la cartographie sémantique	109
3.3.1 Nécessité d'un formalisme dédié	109
3.3.2 Les formalismes de description	110
3.3.3 Présentation de SNDF	116
3.3.4 Synthèse.....	123
3.4 La cartographie sémantique pour l'exploration	125
3.4.1 Positionnement de la problématique	125
3.4.2 Contexte et besoins.....	126
3.4.3 Cartographie de concepts et critères de cartographie	127
3.4.4 EyeTree	132
3.4.5 RadialTree	134
3.4.6 Synthèse.....	135
3.5 MDL : notre langage de cartographie sémantique	137
3.5.1 Nécessité d'un nouveau langage	137
3.5.2 Principe de MDL	138
3.5.3 La structure du langage	140
3.5.4 Exemples	144
3.5.5 Synthèse.....	146
3.6 Notre architecture à agents et MVC distribué	147
3.6.1 Cartographie sémantique orientée service.....	147
3.6.2 Distribution du processus de cartographie.....	148
3.6.3 Synthèse.....	154
3.7 Synthèse des propositions	156

Cette partie présente l'ensemble de nos propositions. Elles constituent un tout cohérent, c'est-à-dire la cartographie sémantique proprement dite.

La première proposition est une **méthode** pour cartographier un espace informationnel selon nos besoins. Elle repose sur la prise en compte de la sémantique du domaine et sur les retours d'expérience. Nous nommons cette méthode la « cartographie sémantique ».

La seconde proposition est un **processus** de cartographie sémantique. La cartographie est alors assimilée à un ensemble d'opérations qui permettent de construire une carte d'un espace informationnel *brut* en passant successivement par un espace informationnel *structuré* puis un espace informationnel *représenté* et enfin un espace informationnel *visualisé*.

Ce processus de cartographie sémantique requiert un certain nombre de modélisations, doit répondre à des critères, et nécessite une architecture pour sa mise en œuvre.

C'est pourquoi, notre troisième proposition est notre **formalisme de description** dédié à la cartographie sémantique, nommé SNDF (Semantic Network Description Formalism). Ce formalisme permet de décrire d'une part les connaissances du domaine en un ensemble d'entités et de relations, et d'autre part de décrire des propriétés de représentation graphique pour ces entités et ces relations.

La quatrième proposition est la définition de **critères pour la cartographie** d'espaces informationnels *structurés* par une ontologie de domaine (comme l'utilisation de techniques « focus+context » avec une géométrie euclidienne) et la spécification de deux nouveaux paradigmes pour explorer un espace informationnel *structuré* : l'EyeTree et le RadialTree.

La cinquième proposition est un **langage** de haut niveau pour construire une carte en décrivant le processus de cartographie sémantique, nommé MDL (Map Description Language).

Enfin, la sixième et dernière proposition est une **architecture** de type système multiagents basée sur une extension du modèle MVC (Model View Controller). Elle permet d'intégrer la cartographie sémantique dans les solutions de gestion de connaissances à l'aide d'un ensemble d'agents logiciels.

3.1 La cartographie sémantique

Notre objectif est de proposer une cartographie capable d'aider les organisations à appréhender leur espace informationnel pour mieux le maîtriser.

Dans notre cas, la cartographie a pour finalité de visualiser (sous forme de carte) un ensemble de données abstraites. Comme nous l'avons décrit précédemment dans ce document, la visualisation constitue un excellent médium entre un grand nombre d'informations et l'esprit. Cependant, il est possible de construire un nombre quasiment infini de cartes pour un même espace informationnel. D'ailleurs ce document ne présentant qu'un panel limité de paradigmes, il est facile d'imaginer la quantité de combinaisons possibles.

Dans ce contexte-là, comment construire une carte ? En effet, il ne suffit pas de produire une carte à partir d'un espace informationnel pour que le résultat soit « efficace ».

« Devant la masse toujours croissante des données, la cartographie offre des possibilités de communiquer synthétiquement et dégager les informations stratégiques utiles aux décideurs. Seulement, comment passer des données brutes à la carte ? » Pierre Mongin (préface de [Poidevin, 1999]).

De plus, une mauvaise carte peut véhiculer une information ambiguë ou erronée et son utilisateur peut se retrouver pénalisé lourdement dans ses prises de décisions.

Pour éviter ce genre de mésaventure, nous avons identifié différents besoins que la cartographie doit satisfaire pour répondre à notre problématique d'appréhension d'un espace informationnel :

Naviguer selon la sémantique du domaine : pour leurs activités, les collaborateurs doivent naviguer dans l'espace informationnel de leur organisation (comme lors d'une recherche d'information). Pour faciliter cette navigation, il est nécessaire que les cartes véhiculent la sémantique du domaine. Si la cartographie respecte cette sémantique, les utilisateurs peuvent alors la comprendre, l'assimiler et l'exploiter et par voie de conséquence, ils peuvent appréhender plus facilement l'ensemble de l'espace informationnel.

Proposer une vision à plusieurs échelles : permettre à l'utilisateur de s'approprier l'information nécessite de lui fournir les moyens pour l'appréhender dans sa globalité mais aussi dans ses particularités. C'est pourquoi, il est nécessaire d'offrir aux collaborateurs simultanément une vision globale et synthétique de l'espace informationnel et une vision particulière et détaillée de ce même espace.

Proposer une carte adaptée à l'utilisateur : tous les collaborateurs n'ont pas la même activité et le même niveau d'expertise. Chaque carte doit donc être adaptée à son utilisateur. De plus, une carte étant une représentation de l'espace informationnel, elle véhicule une signification. Il est alors nécessaire de bien choisir et adapter cette représentation pour maîtriser la signification de la carte qui sera finalement perçue.

L'ensemble de ces besoins est lié à la nécessité pour l'utilisateur de naviguer dans un espace informationnel souvent important. Notre approche pour répondre à cette problématique consiste alors à cartographier l'espace informationnel des organisations.

Ainsi, la cartographie de données abstraites constitue l'axe principal de notre étude. Dans le cadre de la cartographie d'un espace informationnel d'une organisation prenant en compte la sémantique du domaine, nous parlons de cartographie sémantique. Elle devient une activité essentielle à la gestion des connaissances en permettant de tirer parti de toute la richesse des informations de l'organisation.

3.1.1 Cartographie orientée sémantique

Des travaux récents suggèrent que pour réaliser une carte efficace il est nécessaire de respecter le principe fort de « correspondance de représentations » [Chabris & Kosslyn, 2005]. En accord avec ce principe, une représentation graphique est d'autant plus efficace qu'elle correspond à la représentation mentale que se fait l'utilisateur des informations représentées.

Pour un collaborateur, la représentation mentale qu'il se fait d'un espace informationnel « correspond » à la sémantique du domaine¹. Donc pour être en accord avec ce principe de correspondance de représentations, la carte doit respecter la sémantique du domaine.

Notre première proposition définit une cartographie de données abstraites exploitant la sémantique du domaine. Concrètement, cette approche consiste à guider par cette sémantique le choix des techniques à mettre en œuvre pour cartographier un espace informationnel selon les besoins.

Pour comprendre l'importance de la sémantique dans le choix des techniques de cartographie, prenons l'exemple d'un organigramme de collaborateurs.

a. Exemple de cartographie sémantique

Imaginons une entreprise qui possède un ensemble de collaborateurs qui ont au plus un supérieur hiérarchique. L'objectif est d'obtenir une vision globale de l'organisation des collaborateurs sous la forme d'un organigramme.

Si l'on se concentre sur la représentation de l'espace informationnel, nous avons vu qu'il existait de nombreuses possibilités pour construire et répartir les structures visuelles dans une carte.

Voici trois représentations possibles de l'espace informationnel :

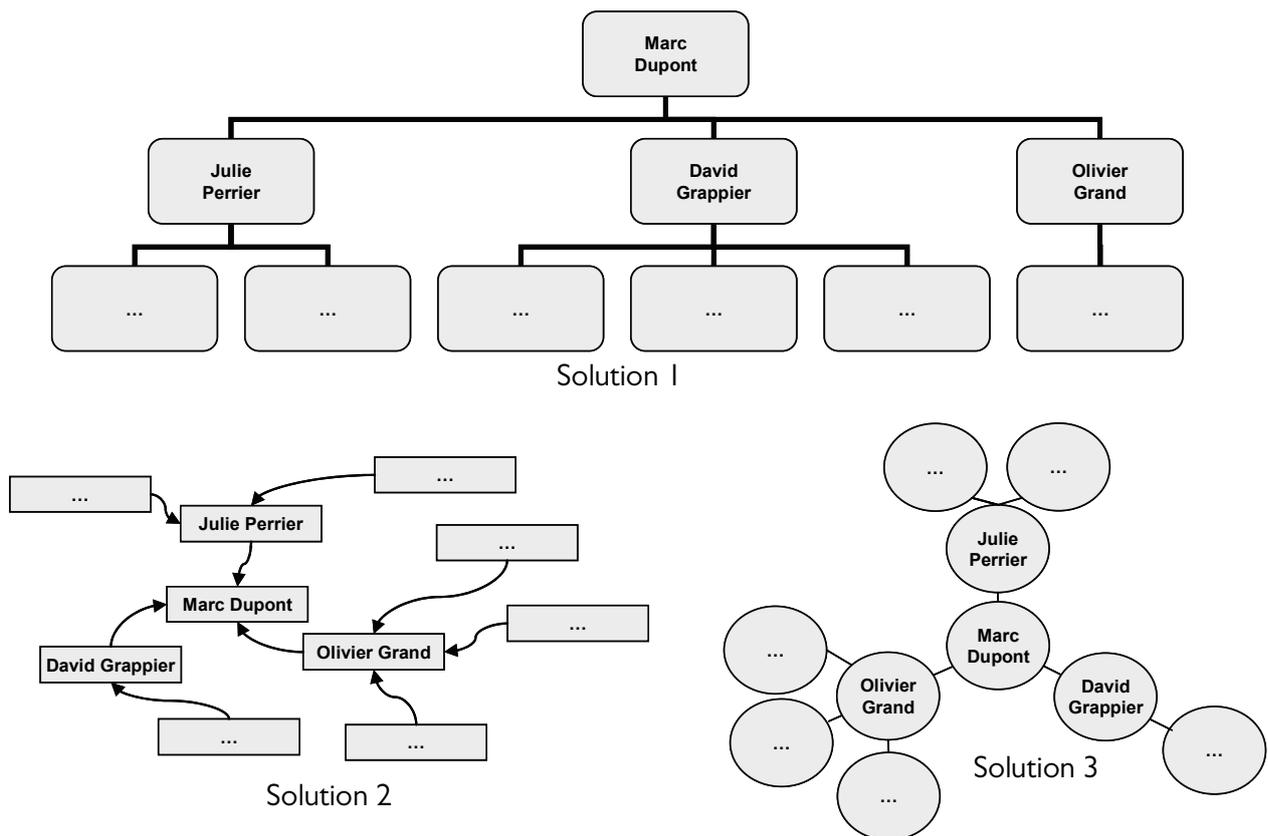


Figure 65 - Exemples de répartitions possibles des représentations des collaborateurs de l'organigramme.

¹ Ensemble des concepts permettant d'appréhender les objets du monde (le domaine).

Dans ce contexte-là, quelle solution choisir ?

À première vue, la difficulté n'est pas d'associer une structure visuelle à chaque collaborateur (par exemple un rectangle ou un cercle) mais bien de les répartir dans l'espace de la carte sachant que cette répartition a un impact important sur la signification perçue.

Généralement pour représenter un organigramme, la première solution est utilisée : chaque collaborateur est placé graphiquement en dessous de son responsable. Intuitivement, cette représentation semble la plus judicieuse, mais pourquoi ?

Pour l'expliquer simplement, prenons la relation de responsabilité qui lie les collaborateurs entre eux. Cette relation a une sémantique de hiérarchie : un responsable est « supérieur » à l'ensemble des collaborateurs qu'il a sous sa responsabilité. C'est pourquoi, pour être cohérent avec le principe de correspondance des représentations, les structures visuelles associées aux collaborateurs doivent respecter cette sémantique. En d'autres termes, chaque représentation d'un collaborateur (sa structure visuelle) doit être placée sous celle de son responsable.

Dans les solutions 2 et 3, les structures visuelles des collaborateurs responsables ne sont pas graphiquement « au dessus » des collaborateurs qu'ils encadrent. La sémantique n'est donc pas graphiquement respectée. Néanmoins ces répartitions peuvent être utilisées pour véhiculer d'autres messages comme le tissu relationnel entre collaborateurs.

En règle générale, la problématique est de choisir la meilleure technique pour exprimer la sémantique du domaine sachant qu'il n'existe aucun modèle universel pour déterminer ce choix.

Pour comprendre la difficulté de l'exercice, attribuons une sémantique un peu différente pour la relation hiérarchique entre collaborateurs ; prenons par exemple, une sémantique orientée « encadrement ». Dans ce cas, une représentation de type Treemaps (voir page 72) peut tout à fait convenir :

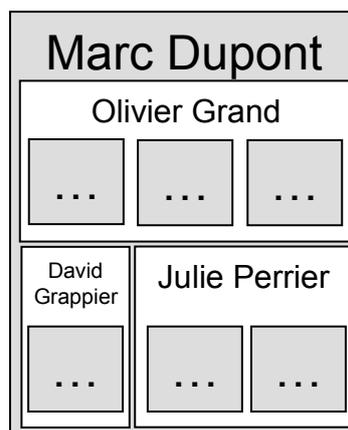


Figure 66 - Représentation de l'organigramme avec une technique de type Treemaps.

Ainsi, la représentation d'un collaborateur (sa structure visuelle) « encadre » celle de ses subordonnées. La carte offre alors une vision ensembliste de l'organisation (et non hiérarchique).

Les paradigmes de représentation et de visualisation déterminent donc le message global véhiculé par la carte.

Bien que la graphique nous aide à déterminer la signification attribuée aux structures visuelles, il n'est pas garanti que tout le monde perçoivent exactement le même message et la même sémantique.

Cette constatation nous permet d'introduire un deuxième aspect de notre démarche de cartographie sémantique : la prise en compte des retours d'expérience des utilisateurs pour adapter les cartes.

3.1.2 Cartographie orientée retours d'expérience

Notre expérience de la cartographie nous a aussi révélé l'importance et la valeur des retours d'expérience pour construire une bonne carte. En effet, même si toutes les cartes sont basées sur la graphique (langage monosémique), une carte ne peut pas être considérée comme définitive et universelle.

L'étude de la graphique et la prise en compte de la sémantique permettent de déterminer les opérations de représentation à mettre en œuvre pour véhiculer une signification particulière. Cependant, dans la réalité et à l'usage, les interprétations peuvent diverger, même légèrement. Surtout, le message perçu par les utilisateurs et leur sentiment vis-à-vis de la carte peuvent fortement varier. C'est d'ailleurs souvent le cas avec des utilisateurs qui ont un niveau différent d'expertise (de connaissance) sur le domaine.

Une mauvaise carte peut induire un mauvais message et les prises de décisions qui en découlent peuvent être préjudiciables.

Depuis l'origine de la cartographie, les cartographes savent qu'une carte se doit d'être esthétique, intuitive et efficace. Malheureusement aucun système ne permet de garantir la satisfaction de ces trois propriétés et encore moins pour tous les individus.

Contrairement aux cartes géographiques, la carte sémantique ne peut être universelle. Elle doit d'une part répondre aux besoins des utilisateurs (décrit précédemment) et d'autre part être adaptée à :

- > **Un espace informationnel** : il est évident que le contenu de la carte détermine sa cartographie ;
- > **Une communauté** : les individus d'une communauté partagent des connaissances utiles pour comprendre et manipuler les informations de la carte. Ne pas exploiter ces connaissances communes pénalise l'efficacité de la carte ;
- > **Un point de vue** : nous ne partageons pas la même vision et interprétation d'un domaine.
- > **Une activité** : une carte est créée pour réaliser une activité (comme la consultation, la modélisation ou la structuration) ;
- > **Un média** : le média (papier, écran, projecteur, etc.) a un impact direct sur le choix des techniques de cartographie.

Face à cette problématique, le plus important pour cartographier un espace informationnel est de toujours se laisser la possibilité de faire évoluer la cartographie selon les retours des utilisateurs. Étant donné qu'une carte sémantique a pour objectif d'aider des utilisateurs, leurs retours d'expérience constituent un critère fiable pour juger de l'efficacité d'une carte. Cette approche est aussi commune à d'autres travaux et plus particulièrement, dans un article qui traite de la difficulté d'évaluer les systèmes de visualisation [Plaisant, 2004].

Nous proposons donc une cartographie basée d'une part, sur la sémantique du domaine pour structurer ses connaissances et pour déterminer les techniques à mettre en œuvre et d'autre part, sur la prise en compte des retours d'expérience pour améliorer et adapter les cartes.

Concrètement, cette approche consiste à recueillir les critiques et les souhaits des utilisateurs. Elles sont l'expression de leurs sentiments et de leurs impressions vis-à-vis de chaque carte. Prendre en compte les retours d'expériences c'est donc tenir compte des critiques (positives comme négatives) et des souhaits des utilisateurs pour adapter la cartographie.

Exprimer des critiques sur les cartes par écrits est jugé difficile par les utilisateurs, c'est pourquoi ils ont rapidement privilégié les échanges de « vive voix ». Ainsi, le téléphone est devenu leur médium principal pour nous communiquer leurs retours d'expérience. Par exemple, voici une retranscription d'un extrait de conversation téléphonique avec un utilisateur qui travaillait avec carte basée sur les arbres hyperboliques :

« J'ai du mal à utiliser les arbres hyperboliques car je suis perdue. Quand je déplace un nœud, les autres se déplacent dans tous les sens surtout aux bords de la carte. Je n'arrive pas à contrôler leurs déplacements. Je finis toujours pas être perdu et à recharger la carte. »

Étant donné que plusieurs témoignages concordaient sur ce point, nous avons donc abandonné le paradigme des arbres hyperboliques au profit d'un paradigme basé sur une géométrie euclidienne. Pour plus de détails sur ce cas et sur notre méthode de prise en compte des retours d'expérience, vous pouvez lire la partie sur la recherche de critères pour l'exploration de bases de connaissances (page 125).

3.1.3 Synthèse

L'espace informationnel détermine le contenu de la carte et les paradigmes (de représentation et de visualisation) utilisés déterminent la perception que peuvent en avoir les utilisateurs. C'est pourquoi, si l'on veut aider les utilisateurs à appréhender leur espace informationnel, il est nécessaire de guider la structuration du domaine et le choix de ces paradigmes par la sémantique du domaine.

Pour répondre aux besoins des utilisateurs, nous proposons donc une cartographie basée d'une part sur la sémantique du domaine et d'autre part sur la prise en compte des retours d'expérience pour améliorer et adapter la carte.

Nous qualifions notre approche de cartographie sémantique que nous définissons comme suit :

La cartographie sémantique est la cartographie de l'espace informationnel d'une organisation basée sur sa sémantique (comprise comme une conceptualisation du domaine).

Cette définition fait référence à la cartographie (générale) qui est définie comme l'ensemble des opérations qui interviennent en vue de l'élaboration de cartes.

3.2 Notre processus de cartographie sémantique

La section précédente présente notre approche de la cartographie qui consiste à exploiter la sémantique du domaine et les retours des utilisateurs pour cartographier un espace informationnel afin de répondre à leurs besoins. Face à ce nouvel enjeu et à la vue des travaux existants, nous avons été amenés à définir un cadre théorique général pour appréhender la cartographie sémantique.

3.2.1 Des données à la carte

La cartographie de données abstraites est un processus qui permet de passer d'un espace informationnel à une carte en appliquant certaines techniques de représentation et de visualisation. Dans la littérature, ces techniques sont assimilées à des opérations.

En accord avec les approches courantes, l'espace informationnel que l'on souhaite cartographier doit tout d'abord être représenté (à l'aide d'opérations de représentation). On obtient alors un ensemble d'objets graphiques nommés « structures visuelles ». Ensuite, cet « espace informationnel représenté » doit être visualisé (à l'aide d'opérations de visualisation) pour obtenir une vue (la carte). Étant donné qu'il existe plusieurs opérations de représentation, il est possible d'obtenir un très grand nombre de représentations différentes d'un même espace informationnel. De même avec les opérations de visualisation, il est possible d'obtenir une grande quantité de visualisations différentes d'un même espace informationnel représenté.

La cartographie de données abstraites est donc assimilable à un ensemble d'opérations appliquées successivement à un espace informationnel que nous qualifions de « brut » (l'espace à cartographier). À chaque opération, l'espace informationnel est traduit vers un autre espace avec au minimum, un espace informationnel représenté et un espace informationnel visualisé (ces différents niveaux sont décrits à la page 52).

Voici un rappel du schéma qui décrit ce processus de cartographie de données abstraites :

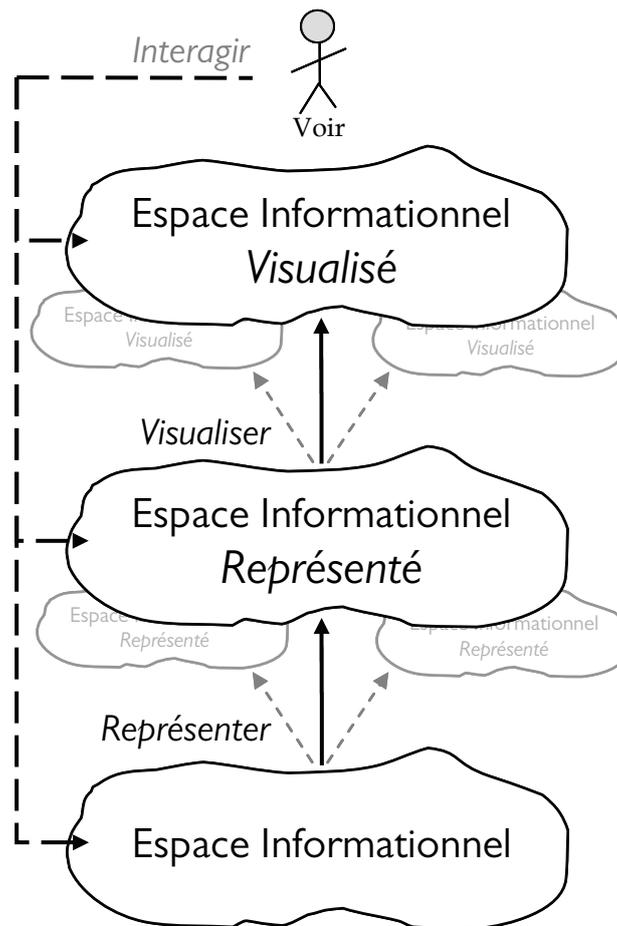


Figure 67 - Les niveaux de la cartographie de données abstraites : de l'information à sa visualisation.

Le modèle de référence de la visualisation d'informations correspond à cette approche en niveaux (voir à la page 47). Les tables de données correspondent à l'espace informationnel. Cependant, ce modèle de référence possède un inconvénient majeur pour répondre à nos besoins : la structuration des informations à cartographier en tables de données.

Pour rappel, le cœur de notre approche repose sur la prise en compte de la sémantique du domaine pour guider la cartographie. Cet objectif implique d'explicitier cette sémantique (à l'aide d'une modélisation). Seulement avec les tables de données, la sémantique est explicitée par le nom des colonnes (nommées « métadonnées »). Par conséquent, elle est limitée et elle n'est pas facilement manipulable. Sans ou avec peu de sémantique, les données à cartographier restent des données « sans sens », c'est pourquoi nous nommons cette approche : « des données à la carte ».

Notre objectif est d'associer la sémantique du domaine à la cartographie ce qui revient à passer des données aux connaissances c'est pourquoi, nous nommons notre approche : « des connaissances à la carte ».

Dans ce dessein, nous proposons un nouveau processus dédié à la cartographie sémantique.

3.2.2 Des connaissances à la carte

Le cœur de notre approche repose sur la prise en compte de la sémantique du domaine pour guider la cartographie selon les besoins des utilisateurs. Cette approche implique tout d'abord d'explicitement cette sémantique.

a. Explicitation de la sémantique

Dans le cadre de nos travaux, nous définissons la sémantique du domaine comme l'ensemble des concepts permettant d'appréhender les objets du monde (le domaine). Par extension et en d'autres termes, la sémantique d'un domaine correspond à l'ensemble de ses concepts.

Pour expliciter la sémantique, nécessaire à notre approche, il faut donc décrire les concepts du domaine ainsi que leurs relations [Roche, 2005]. Cette description définit alors une modélisation des connaissances du domaine.

Pour rappel, nous considérons les connaissances d'une organisation comme l'ensemble de son patrimoine intellectuel caractérisé par ses connaissances explicites (l'espace informationnel contenant les documents, projets, retours d'expérience, etc....) et les connaissances tacites détenues par les collaborateurs et les experts (les expériences, le savoir-faire, etc. ...).

L'explicitation de la sémantique se fait par la modélisation des connaissances conceptuelles du domaine. La première étape dans le processus de cartographie sémantique consiste donc à **modéliser** le domaine.

b. Plusieurs modélisations (et structurations) d'un même espace informationnel

Chaque collaborateur d'une organisation possède sa propre vision de son domaine et par conséquent, chaque collaborateur possède aussi ses propres connaissances sur ce domaine. Il est donc nécessaire de proposer un processus de cartographie capable de supporter plusieurs modélisations du domaine correspondant à autant de points de vue théoriques différents.

Pour illustrer la possibilité de concevoir plusieurs modélisations d'un même monde, prenons l'exemple de la lumière. Dans un premier cas, la lumière peut être modélisée selon un point de vue corpusculaire : « la lumière est un flux qui contient des particules ». Dans un second cas, la lumière peut être modélisée selon un point de vue ondulatoire : « La lumière est une onde ». Ces deux points de vue théoriques produisent donc bien deux modélisations différentes.

Pour permettre aux utilisateurs de naviguer dans l'espace informationnel selon cette sémantique, il est alors nécessaire de restructurer l'espace informationnel de départ (*brut*) selon la modélisation des connaissances du domaine. Le résultat est alors un nouvel espace que nous qualifions d'espace informationnel *structuré*.

La deuxième étape dans le processus de cartographie sémantique consiste donc à **structurer** l'espace informationnel à cartographier. Nous nommons le résultat, « l'espace informationnel *structuré* ».

c. Plusieurs représentations d'un même espace informationnel *structuré*

Une fois structuré, l'espace informationnel doit être représenté graphiquement. C'est-à-dire, qu'il est nécessaire de créer un nouvel espace constitué de structures visuelles. Comme nous l'avons vu précédemment, il existe de nombreuses techniques de représentations qui peuvent être combinées pour réaliser une représentation.

La représentation de l'espace informationnel doit permettre de répondre à notre besoin d'adapter les cartes à l'utilisateur avec une navigation basée sur la sémantique du domaine.

Chaque utilisateur possède un profil particulier (avec une activité et un niveau d'expertise donné). Une représentation unique ne peut donc convenir. Il est donc nécessaire de proposer un processus de cartographie capable de supporter plusieurs représentations d'un même espace informationnel *structuré* correspondant à autant de cartes différentes. La modélisation permet alors de sélectionner les opérations de représentations qui respectent la sémantique.

L'application d'opérations de représentations permet de construire des espaces de nature graphique qui seront perçus par les utilisateurs au travers de leur visualisation. Par conséquent, chaque espace doit (en plus de respecter la sémantique) répondre à des critères esthétiques. En effet, de nombreux travaux dédiés à la représentation de graphes [Purchase, 1997; Ware et al., 2002], présentent le caractère déterminant de l'esthétique dans la compréhension d'un ensemble de données par des utilisateurs. Par conséquent, le choix des opérations doit être guidé par la sémantique du domaine mais aussi par le respect de propriétés esthétiques.

La troisième étape dans le processus de cartographie sémantique consiste donc à **représenter** l'espace informationnel *structuré*. Nous nommons le résultat, « l'espace informationnel *représenté* ».

d. Plusieurs visualisations d'un même espace informationnel *représenté*

Pour être perceptible par l'utilisateur, un espace informationnel *représenté* (par un ensemble de structures visuelles) doit être visualisé. Comme pour la représentation, cette opération de visualisation doit permettre de répondre à notre besoin d'adapter les cartes à l'utilisateur.

En effet, chaque utilisateur possède un support particulier pour afficher la carte (écran, projecteur, papier, etc.) avec des dimensions toutes aussi particulières. Une visualisation unique ne peut donc pas convenir. Il est alors nécessaire de proposer un processus de cartographie capable de supporter plusieurs visualisations d'un même espace informationnel *représenté* correspondant à autant de cartes différentes.

La quatrième étape dans le processus de cartographie sémantique consiste donc à **visualiser** l'espace informationnel *représenté*. Nous nommons le résultat, « l'espace informationnel *visualisé* » ou « vue » ou bien encore, « la carte ».

e. Points de vue simultanés & interactions

Lors de la navigation de l'espace informationnel, le point de vue des utilisateurs peut changer tout comme leur activité et leur niveau d'expertise. De plus, il est souvent nécessaire qu'un même utilisateur appréhende simultanément un espace informationnel selon plusieurs points de vue théoriques et pour différentes activités.

Par conséquent, il est nécessaire que le processus de cartographie sémantique ne soit pas linéaire et unique. L'utilisateur doit pouvoir appréhender son espace informationnel selon différents points de vue correspondant à autant de cartes différentes.

Par exemple, prenons le domaine de la cartographie de données abstraites. Imaginons des experts qui souhaitent modéliser graphiquement ce domaine en structurant tous ses concepts. On imagine vite le nombre de concepts à représenter. Étant donné que cette cartographie est pluridisciplinaire, elle met en jeu plusieurs domaines de compétences. Une approche de modélisation pourrait alors consister à représenter chaque domaine dans une vue dédiée ainsi que tous les concepts qui s'y rapportent. Les experts pourraient alors se concentrer sur leur

domaine. De même, une vue pourrait aussi être utilisée pour visualiser tous les concepts ou entités liés par une relation particulière.

De plus, l'utilisateur doit pouvoir adapter la cartographie à ses besoins. Concrètement, il doit modifier ou changer les différentes opérations de chacune des cartes pour répondre à ses besoins. Dans ce dessein, notre processus de cartographie doit mettre en œuvre des techniques d'interaction pour adapter les différents espaces.

La cartographie d'un même espace informationnel selon différentes cartes adaptables par l'utilisateur permet de prendre en compte les retours d'expérience. À tout moment l'utilisateur doit avoir une carte adaptée.

La cinquième étape dans le processus de cartographie sémantique consiste donc à donner les moyens à l'utilisateur pour **adapter** sa carte (l'espace informationnel *visualisé*). À chaque interaction, la cartographie est mise à jour en réitérant les étapes nécessaires. Le processus de cartographie sémantique est donc cyclique.

3.2.3 Description du processus

Notre approche permet de cartographier sémantiquement en une ou plusieurs cartes un même espace informationnel.

La construction de chaque carte peut être assimilée à un ensemble d'opérations permettant de passer successivement par :

1. Un espace informationnel **brut** : il s'agit de l'espace initial que l'on souhaite cartographier.
2. Un espace informationnel **structuré** : cet espace est obtenu par une opération de structuration de l'espace informationnel basée sur une modélisation construite au préalable.
3. Un espace informationnel **représenté** : cet espace est obtenu par des opérations de représentation de l'espace informationnel *structuré* (le choix des opérations est en accord avec la modélisation).
4. Un espace informationnel **visualisé** : cet espace est obtenu par des opérations de visualisation de l'espace informationnel *représenté* (le choix des opérations est, lui aussi, déterminé en accord avec la modélisation).
5. Une carte **adaptée** : l'utilisateur adapte ses cartes en interagissant avec les différents espaces de la cartographie.

Voici donc un schéma qui illustre notre processus de cartographie sémantique :

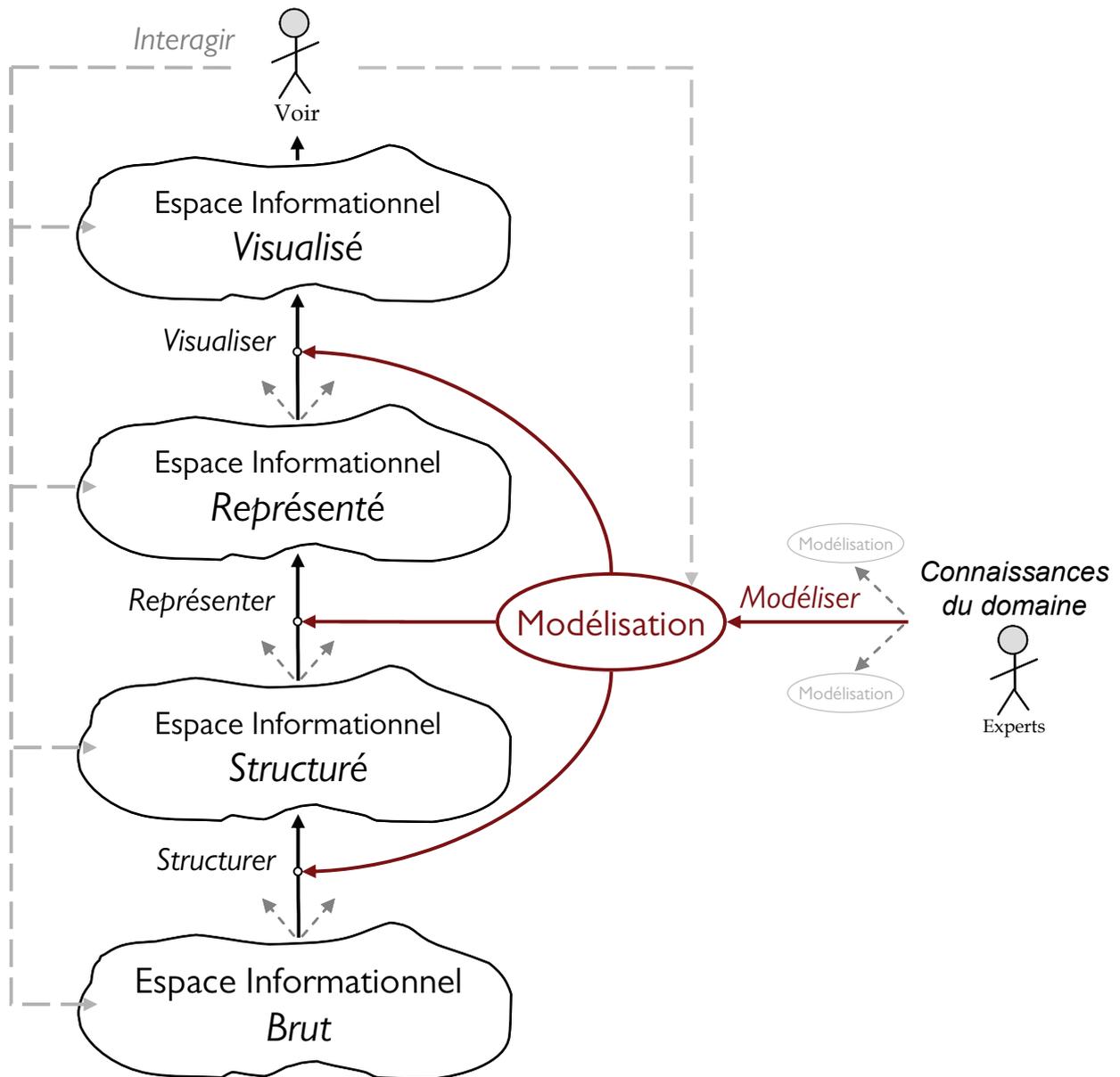


Figure 68 - Notre processus de cartographie sémantique.

Le processus de cartographie sémantique n'est pas un processus linéaire. Tout d'abord, pour un espace informationnel donné, notre processus permet d'en construire plusieurs structurations (selon plusieurs modélisations). Ensuite, pour un espace *structuré* donné, notre processus permet d'en construire plusieurs représentations. Et enfin, pour un espace *représenté* donné, il est possible d'en avoir plusieurs visualisations.

Voici une illustration avec un exemple simple de cartographie d'organigramme de collaborateurs selon différentes opérations :

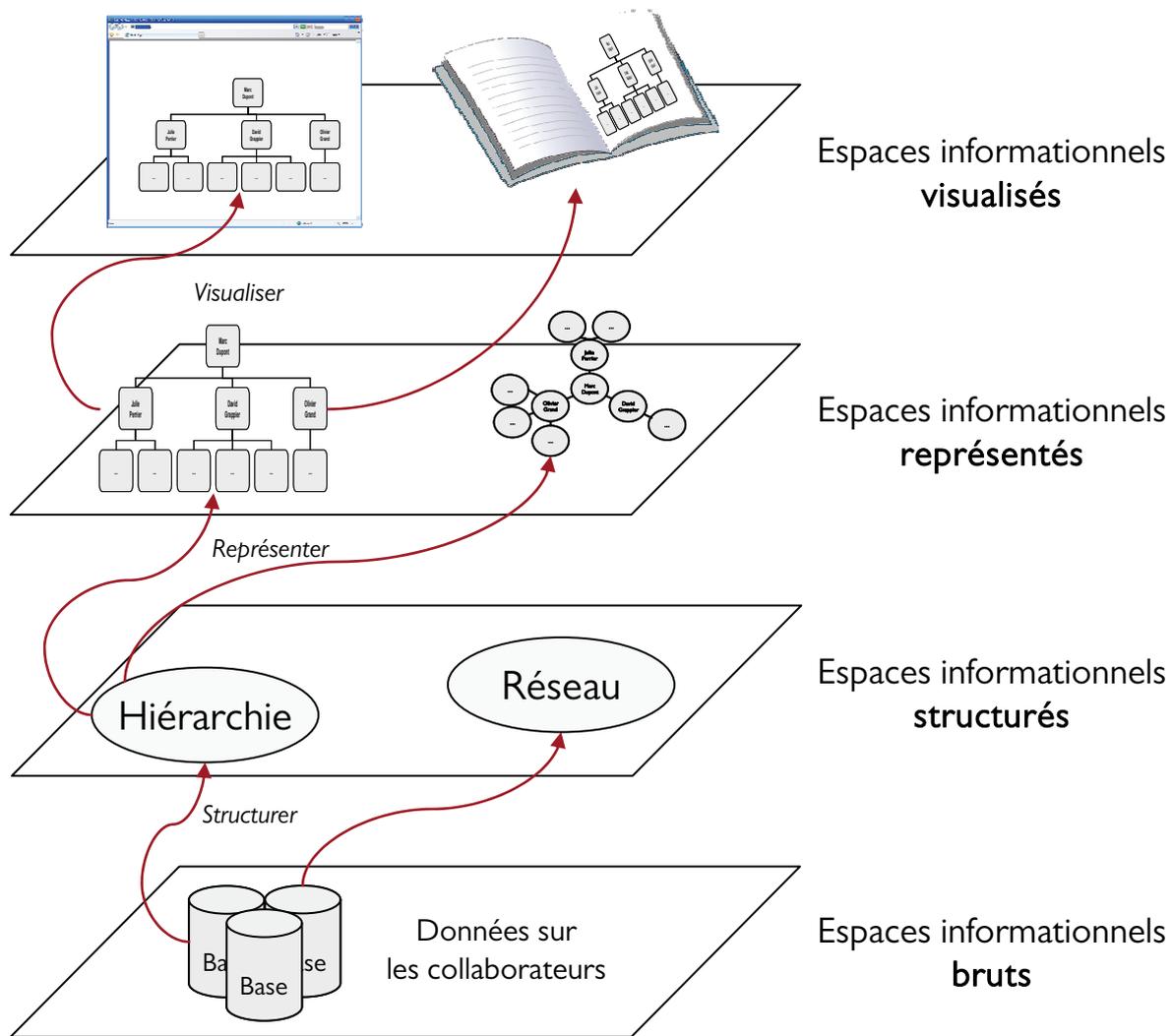


Figure 69 - Illustration de notre processus avec un organigramme.

Notre proposition de processus assimile la cartographie sémantique à une succession de changements d'espaces effectués par des opérations.

a. Les espaces

Notre processus introduit la notion abstraite d'espace pour décrire le contenu de chaque étape de la cartographie. Cette notion abstraite permet de décrire symboliquement les espaces selon la nature de leurs éléments (abstraite, graphique ou visuelle). Ainsi dans notre processus, chaque espace possède sa propre théorie pour exprimer les éléments qu'il contient.

Remarque : cette approche symbolique permet pour décrire chaque espace, d'utiliser des théories issues de l'intelligence artificielle (comme la théorie des relations, la logique, les réseaux sémantiques, etc.) [Gordon, 2000; Sowa, 2000].

Dans notre processus, nous distinguons quatre natures d'espaces : *brut*, *structuré*, *représenté* et *visualisé*. Pour construire une carte, il est nécessaire de passer par au moins un espace de chaque nature. En effet, il est impossible de passer d'un espace informationnel à une carte sans passer par une représentation et une visualisation. Chaque espace est un prérequis pour construire l'espace suivant.

b. Les opérations

Les opérations sont les transformations appliquées aux espaces. D'un point de vue opérationnel, le processus de cartographie sémantique est assimilable à une succession d'opérations.

Chaque opération porte sur un espace et produit un nouvel espace. Par exemple, une opération de type zoom permet de passer d'une vue (espace informationnel *visualisé*) à une autre vue. De même, une opération de distorsion permet de passer d'un espace de représentation (espace informationnel *représenté*) à un espace de vue. Pour assurer la cohérence du processus, il est nécessaire de s'assurer qu'une opération sur un espace produit un nouvel espace de nature équivalente ou de nature immédiatement inférieure ou immédiatement supérieure.

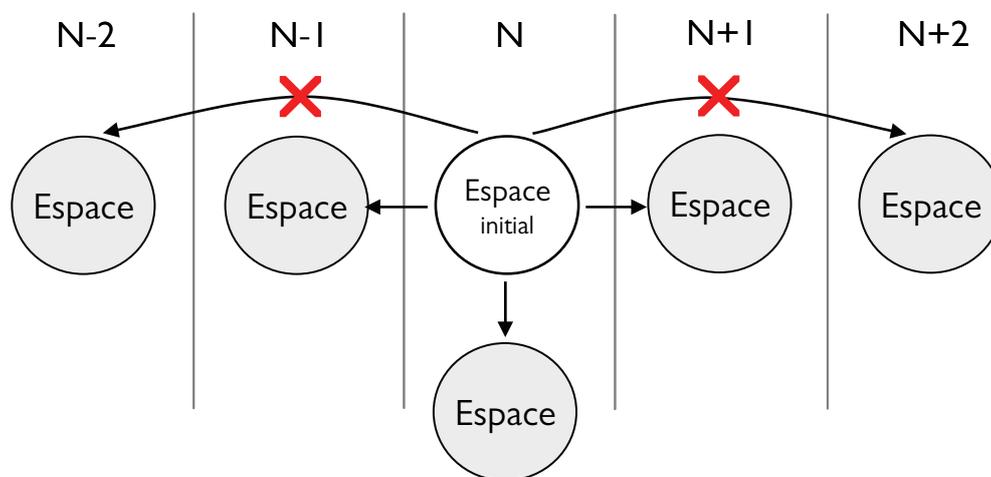


Figure 70 - Changements de nature d'espace possibles lors d'une opération.

Chaque opération porte sur un espace (dit initial) et produit un nouvel espace. Elle peut aussi accepter en entrée un ensemble de paramètres utiles aux traitements associés. Par exemple, une opération de type « zoom » porte sur un espace *visualisé* et elle prend en entrée un vecteur pour paramétrer l'homothétie. De même, une opération de « filtrage » prend en paramètre le critère du filtre.

Cette approche permet de classer les opérations en fonction de la nature de l'espace initial et sur la nature de l'espace produit. Nous avons donc étendu des résultats existants [Card et al., 1999d; Chi & Riedl, 1998; Tweedie, 1997] pour fournir ce classement des principales opérations possibles entre les espaces :

- > Un espace informationnel *brut* vers ...
 - ... un autre espace informationnel *brut* :
 - Nettoyage des données, traitement des données manquantes, filtrage des données ;
 - ... un espace informationnel *structuré* :

- Modélisation, structuration, clustering¹, classification², extraction des variables quantitatives, vectorisation ;
- > Un espace informationnel *structuré* vers ...
 - ... un autre espace informationnel *structuré* :
 - Classification, clustering, recherche de vecteurs, intersection de vecteurs, filtrage des éléments, normalisation ;
 - ... un espace informationnel *représenté* :
 - Codage (création de structures visuelle), répartition dans l'espace de la carte ;
 - ... un espace informationnel *brut* :
 - Construction de valeurs dérivées ;
- > Un espace informationnel *représenté* vers ...
 - ... un autre espace informationnel *représenté* :
 - Filtrage des structures visuelles (selon des propriétés graphiques), fusion, sélection ;
 - ... un espace informationnel *visualisé* :
 - Visualisation, déformation ;
 - ... un espace informationnel *structuré* :
 - Décodage ;
- > Un espace informationnel *visualisé* vers ...
 - ... un autre espace informationnel *visualisé* :
 - Changement de point de vue
 - Rotation, homothétie, défilement, zoom
 - Colorisation, filtrage, permutation d'axes, changement de niveau de détails, estompage, cacher sous arbre, pivoter sous arbre ;

Cette liste d'opération n'est pas exhaustive. Elle a seulement pour objectif de donner une idée des opérations possibles entre les espaces.

3.2.4 Synthèse

Le cœur de notre approche repose donc sur la prise en compte de la sémantique du domaine pour guider la cartographie selon les besoins des utilisateurs. Cet objectif implique d'explicitier cette sémantique à l'aide d'une modélisation qui décrit les concepts du domaine ainsi que leurs relations et ensuite de structurer l'espace informationnel selon cette modélisation.

¹ Une opération de clustering permet d'identifier un modèle dans les données (des clusters) et d'organiser ces données selon ce modèle. Le modèle n'est pas connu à priori.

² Une opération de classification permet d'organiser les données selon un ensemble de classes définies à priori.

Notre proposition de processus permet de prendre en compte la sémantique du domaine (au travers de sa modélisation) tout au long de la cartographie. De plus, ce processus permet d'assimiler la cartographie à un ensemble d'opérations permettant de passer successivement par : un espace informationnel *brut*, un espace informationnel *structuré*, un espace informationnel *représenté* et enfin, un espace informationnel *visualisé*.

3.3 SNDF : notre formalisme de description pour la cartographie sémantique

Le cœur de notre approche repose sur la prise en compte de la sémantique du domaine pour guider la cartographie. Cette approche nécessite deux phases importantes :

1. Expliciter la sémantique : il faut décrire les concepts du domaine ainsi que leurs relations. Cette description définit une modélisation du domaine.
2. Structurer l'espace informationnel : une fois la modélisation construite, il est nécessaire de structurer l'espace à cartographier selon cette modélisation.

La modélisation joue le rôle d'une lunette d'observation, au travers de laquelle, l'utilisateur va percevoir son espace informationnel (les objets sont décrits selon les concepts du domaine). Cette métaphore illustre l'importance de la modélisation dans la cartographie. La perception de la carte par l'utilisateur est fonction *in fine* de la modélisation utilisée pour structurer l'espace.

La problématique est alors de savoir quel formalisme de description utiliser pour décrire une modélisation et l'espace informationnel *structuré* associé.

3.3.1 Nécessité d'un formalisme dédié

Notre problématique nécessite un formalisme dédié car nous avons besoin de décrire des connaissances particulières. En effet, la cartographie sémantique nécessite d'une part de décrire les connaissances du domaine en un ensemble d'entités et de relations entre ces entités et d'autre part elle a besoin de décrire des propriétés de représentation propres aux éléments de l'espace.

Ces deux types de connaissances à décrire sont présentés dans les deux sections suivantes.

a. Décrire les connaissances du domaine

Notre approche de la cartographie sémantique nécessite un formalisme pour décrire des modélisations et des espaces informationnels *structurés*. Dans ce dessein, il doit permettre de décrire les concepts du domaine (la modélisation) et les éléments de l'espace informationnel selon ces concepts (espace informationnel *structuré*).

Toute description d'un ensemble d'informations (ou espace) repose sur un formalisme porteur d'une théorie qui conditionne la définition de concepts. Par exemple en génie logiciel, pour définir un modèle conceptuel des données (MCD), la théorie utilisée est de type entité-association ; une fois le MCD construit, il fournit les concepts (les tables) pour construire la base de données (structuration de l'espace informationnel ; les objets deviennent des tuples). Dans notre cas, les concepts définis selon la théorie doivent décrire la modélisation du domaine et ils sont utilisés pour décrire les objets de l'espace informationnel.

De plus, notre formalisme doit être adapté à notre processus. C'est-à-dire qu'il doit permettre de manipuler plusieurs modélisations simultanément d'un même espace informationnel. Chaque modélisation permet d'exprimer un « point de vue théorique » sur un domaine. Par conséquent, un même « objet » (instance), s'il est lié à un seul concept par la relation d'instanciation dans une modélisation donnée (ou du moins devrait l'être), peut être

lié à un autre concept par la même relation dans une autre modélisation. Le même objet peut être appréhendé (conceptualisé) selon plusieurs points de vue théoriques différents. Notre formalisme doit donc permettre de partager les objets entre les modélisations.

Remarque : la théorie du formalisme est parfois aussi nommée « modèle ». Cependant, pour éviter toute confusion avec la modélisation de la cartographie, nous préférons utiliser le terme « théorie ».

b. Décrire les propriétés de représentation

Une fois structuré par la modélisation, l'espace doit être représenté puis visualisé pour produire des cartes. Les éléments de l'espace sont alors visualisés dans différentes cartes et l'utilisateur peut interagir avec leur représentation (les structures visuelles).

La difficulté avec la cartographie de données abstraites c'est que les données ne possèdent pas de représentation intrinsèque. La construction de l'espace de représentation permet alors d'associer arbitrairement des structures visuelles aux éléments de l'espace *structuré* (les descriptions d'espaces informationnels *structurés* vont servir à la construction d'espaces informationnels *représentés* puis *visualisés*). Pour assurer une cohérence de représentation et d'interaction entre les cartes, la cartographie attribue généralement aux éléments de l'espace des propriétés communes à leurs structures visuelles dans toutes les cartes. Cette démarche revient à définir une représentation intrinsèque aux éléments de l'espace informationnel. Par conséquent, notre formalisme doit permettre de décrire ces propriétés de représentation pour chaque élément de l'espace informationnel *structuré*.

Pour conclure sur les besoins, notre problématique est de proposer un formalisme pour d'une part décrire les concepts du domaine et structurer l'espace selon ces concepts et d'autre part, pour définir des propriétés graphiques pour les différents éléments de l'espace (concepts, objets, relations, etc.).

Face à cette problématique, nous avons commencé par étudier les différents formalismes de description existants (principalement les langages de représentation des connaissances issues de l'intelligence artificielle).

3.3.2 Les formalismes de description

La description d'un ensemble d'informations se fait dans un formalisme (ou langage) donné. Chaque formalisme repose sur une théorie qui définit ses principes de description. Le choix d'un formalisme et plus particulièrement sa théorie dépend directement de la nature des connaissances à décrire.

a. La typologie des connaissances

Il existe différents types de connaissances selon leur nature. À chaque type de connaissances sont associées une ou plusieurs théories et langages qui permettent de les décrire.

Nous pouvons distinguer quatre types correspondant à autant de niveaux de connaissances, chaque niveau reposant sur le ou les niveaux précédents :

- > **Niveau ontologique** : à ce niveau, les connaissances portent sur la définition des concepts et des relations nécessaires à la description des objets du monde. Ces connaissances sont toujours vraies quel que soit l'état du monde.

- > **Niveau des faits** : à ce niveau, les connaissances décrivent un état du monde c'est-à-dire des faits portant sur les objets (attributs valués) et des relations. Contrairement au niveau précédent, ces connaissances peuvent être vraies ou fausses selon l'état du monde à un moment donné.
- > **Niveau des raisonnements** : à ce niveau, les connaissances inférentielles (règles d'inférences) permettent de déduire de nouveaux faits.
- > **Niveau des heuristiques** : les heuristiques guident le raisonnement en gérant par exemple les règles d'inférences potentiellement applicables à un moment donné.

La cartographie sémantique d'un espace informationnel permet de produire des cartes contenant des représentations d'entités de l'espace et de leurs relations. La modélisation sous-jacente ainsi que l'espace *structuré* associés à la carte, doivent décrire des entités et des relations entre ces entités. Notre formalisme de description (et sa théorie) doit donc appartenir aux deux premiers niveaux de notre typologie des connaissances.

Par conséquent, nous pouvons écarter les théories correspondant aux raisonnements et aux heuristiques. La carte peut être perçue comme une représentation descriptive et déclarative (ici graphique) du monde.

Remarque : bien que nous écartons les deux derniers niveaux, nous pouvons néanmoins cartographier un espace informationnel correspondant à un raisonnement (ou un heuristique) à condition que sa représentation le permette, par exemple sous la forme de graphes d'états.

Étant donné que nous écartons les deux derniers niveaux de la typologie, le lecteur peut s'interroger sur la cartographie d'un espace informationnel contenant des connaissances de raisonnements ou d'heuristiques. C'est par exemple le cas avec des représentations graphiques de raisonnements comme des graphes d'états ; l'espace informationnel *structuré* (selon la modélisation) associé à chaque carte contient des descriptions de raisonnement ou d'heuristiques. Dans ce cas, la description appartiendrait aussi aux deux premiers niveaux de notre typologie, niveau des descriptions des objets du monde (raisonnement ou heuristique) et de leurs relations.

La partie suivante présente les principales théories et leur formalisme qui permettent de décrire les objets du monde et leurs relations.

b. La description des objets du monde et leurs relations

Pour décrire les objets du monde et leurs relations, le domaine de l'intelligence artificielle (IA) offre plusieurs formalismes de description. Chacun d'eux repose sur une théorie propre.

Remarque : les formalismes de description sont généralement nommés « formalisme de représentations ». Cependant, bien que toute description soit nécessairement représentée et pour éviter toute confusion avec la notion de représentation de la cartographie, nous préférons parler de description plutôt que de représentation.

Pour décrire les objets du monde et leurs relations, il existe trois types de formalisme selon leur « degré de formalisation » :

- > Les formalismes « formels » : ces formalismes reposent sur des concepts théoriques forts (comme des opérations formelles de réécriture).
- > Les formalismes « semi-formels » : ces formalismes reposent sur des concepts théoriques plus faibles mais plus proches d'une représentation naturelle des choses.

- > Les formalismes « informels » : ces formalismes correspondent aux langages naturels.

Étant donné que notre objectif est de décrire des ensembles d'informations manipulables par une machine, nous pouvons écarter les formalismes informels.

Les sections suivantes présentent les principaux formalismes (formels puis semi-formels) de description utilisés en intelligence artificielle pour décrire les objets du monde et leurs relations.

i. Les formalismes « formels »

Les logiques

La logique est à l'origine une réflexion sur la cohérence du discours. Elle trouve ces origines à l'Antiquité grecque. Jusqu'au Moyen-âge, elle constitue une des plus grandes disciplines de la philosophie. La fonction première de la logique est d'assurer la cohérence dans les discours (logos). Par la suite elle est devenue un outil assurant la cohérence dans les raisonnements.

Depuis le XIXème siècle puis au XXème, elle est devenue une discipline mathématique et informatique. Dès lors, la logique se diversifie et de nombreux langages basés sur les principes de la logique voient le jour. On parle alors de logiques formelles ou de langages logiques ou tout simplement de logiques.

Les logiques formelles permettent de décrire des objets et leurs relations [Barr & Feigenbaum, 1982 (p.154)]. Une des raisons qui permettent d'expliquer le succès des descriptions basées sur les logiques, est que la déduction de nouveaux faits à partir d'anciens peut être automatisée. En effet en logique, il existe un ensemble de règles nommées règles d'inférences pour lesquelles les faits connus comme vrais peuvent être utilisés pour déduire d'autres faits alors nécessairement vrais.

Pour la description de connaissances, les deux logiques principales sont : le calcul des prédicats (ou logique des prédicats) et les logiques des descriptions.

Le calcul des prédicats

Le calcul des prédicats est une logique formelle proposée par Frege (1879) puis Russell (1910-1913). Elle est aussi parfois nommée calcul des prédicats du premier ordre ou logique du premier ordre.

Le calcul des prédicats a pour fonction initiale de formaliser des raisonnements logiques. Sa spécificité réside dans l'introduction de plusieurs éléments :

- > un ensemble de symboles désignant des variables,
- > un ensemble de symboles désignant des fonctions,
- > un ensemble de symboles désignant des prédicats,
- > des connecteurs logiques et,
- > deux quantificateurs (représentés par les symboles \forall et \exists),
- > des règles d'inférences.

Remarque : on parle de logique du premier ordre car on ne peut quantifier que les variables.

Les logiques des descriptions

Tout comme le calcul des prédicats, les logiques des descriptions sont une logique formelle particulière [Baader *et al.*, 2003; Napoli, 1997]. Elles forment une famille de langages de description de connaissances.

Les logiques des descriptions permettent de représenter les connaissances d'un domaine selon :

- > Des concepts : ils correspondent à des classes d'individus et ils sont organisés en hiérarchies selon une relation de subsomption.
- > Des rôles : ils correspondent aux relations entre les concepts et ils sont aussi organisés en hiérarchies selon une relation de subsomption.
- > Des individus : ils correspondent aux éléments d'un univers donné (le domaine).

Une sémantique est associée aux descriptions par l'intermédiaire d'une fonction d'interprétation.

Ces derniers temps, les logiques de description bénéficient d'un regain d'intérêt par l'intermédiaire des travaux sur le Web Sémantique et plus particulièrement avec OWL¹ (Resource Description Framework). OWL est un format d'échange basé sur les logiques des descriptions et il constitue la base du Web sémantique.

Les ontologies

Les ontologies sont associées au premier niveau de notre typologie des connaissances : descriptions des objets du monde. Bien que très anciennes, les ontologies font l'objet à l'heure actuelle d'un essor et d'un engouement sans précédent. Cette popularité grandissante n'a malheureusement toujours pas permis d'arriver à un consensus sur le sens même des ontologies [Guarino & Giaretta, 1995; Roche, 2003].

L'ontologie, étymologiquement « ontos » (être) et « logos » (science, langage), s'intéresse à la « science de l'être en tant qu'être ». Elle est donc du ressort de la métaphysique (le terme lui-même apparaît tardivement en 1692, emprunté au latin scientifique *ontologia* (1646)). Dans son acception actuelle et pour l'étude qui nous intéresse ici, le terme d'« ontologie » a une signification plus large, il désigne l'ensemble des connaissances relatives à un domaine : objets, concepts, relations et propriétés [Roche, 2005].

Les ontologies ont une visée normative, mais *quid* de la définition du terme « ontologie » lui-même ? L'omniprésence des ontologies en est peut-être la principale raison.

Nous pouvons néanmoins citer quelques définitions, selon le point de vue considéré :

Celui de la Connaissance :

Les ontologies sont une branche de la gestion des connaissances qui s'intéresse à la représentation et à l'organisation des concepts d'un domaine. La définition la plus citée est celle de Thomas R. Gruber :

« An ontology is a specification of a conceptualization. [...] That is, an ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents. This definition is consistent with the usage of ontology as set-of-concept-definitions, but more general. » [Gruber, 1993]

¹ World Wide Web Consortium - <http://www.w3.org/2004/OWL>

Celui d'un Vocabulaire Commun :

« An [explicit] ontology may take a variety of forms, but necessarily it will include a vocabulary of terms and some specification of their meaning (i.e. definitions). » [Uschold & Grüninger, 1996]

Voici la définition de John F. Sowa :

« The subject of ontology is the study of the categories of things that exist or may exist in some domain. The product of such a study, called an ontology, is a catalogue of the types of things that are assumed to exist in a domain of interest D from the perspective of a person who uses a language L for the purpose of talking about D. The types in the ontology represent the predicates, word senses, or concept and relation types of the language L when used to discuss topics in the domain D. » [Sowa, 2000]

Pour conclure, voici la définition que nous avons adoptée :

« Une ontologie est une conceptualisation d'un domaine à laquelle sont associés un ou plusieurs vocabulaires de termes. Les concepts se structurent en un système et participent à la signification des termes. Une ontologie est définie pour un objectif donné et exprime un point de vue partagé par une communauté. Une ontologie s'exprime dans un langage (représentation) qui repose sur une théorie (sémantique) garante des propriétés de l'ontologie en termes de consensus, cohérence, réutilisation et partage. » [Roche, 2005]

L'ontologie est une théorie qui permet de décrire les objets du monde (leur essence et leur structure) indépendamment de leurs manifestations particulières.

Pour permettre l'échange et le partage d'ontologie, le W3C¹ a spécifié un langage de description d'ontologies nommé OWL (Ontology Web Language). Il est basé sur les logiques des descriptions et il permet de décrire des ontologies comme étant des terminologies :

« An ontology defines the terms used to describe and represent an area of knowledge. »²

Les ontologies OWL sont constituées de concepts et de propriétés (aussi appelés rôles en logiques de description). En OWL, un domaine se compose d'instance de concepts.

ii. Les formalismes semi-formels

Les réseaux sémantiques

Les réseaux sémantiques ont été proposés par Quillian en 1968 pour représenter explicitement un modèle psychologique de la mémoire associative humaine [Barr & Feigenbaum, 1982 (p.156); Sowa, 2000].

Un réseau sémantique est composé d'une part, de nœuds qui peuvent représenter indifféremment des objets, des concepts ou des événements et d'autre part, de liens (arcs orientés étiquetés) entre les nœuds qui représentent leurs relations. Chaque réseau

¹ World Wide Web Consortium - <http://www.w3.org/2004/OWL>

² <http://www.w3.org/TR/webont-req/#onto-def>

sémantique est donc une structure de graphe dédiée à la description d'objets et de leurs relations binaires.

De plus, ils peuvent aussi servir de représentation graphique de prédicats binaires de la logique des prédicats [Nilsson, 1980 (p.370)]. Les symboles de variables, fonctions et prédicats sont alors représentés par un réseau de nœuds : les variables et les fonctions sont représentées par des nœuds et les prédicats binaires par des liens.

Les systèmes à base de schémas

Les systèmes à base de schémas ont été introduits en 1932 par le psychologue Frederic Bartlett et mis en œuvre en 1975 par Marvin Minsky [Minsky, 1975].

L'idée à l'origine de ces systèmes est d'imiter le comportement humain face à un problème donné : on recherche dans notre mémoire un modèle que l'on cherche à adapter en changeant un certain nombre de détails. Notre comportement n'est donc pas toujours inférentiel.

Dans le cadre de cette étude, nous nous intéressons aux systèmes à base de schémas pour leur finalité de représentation des connaissances.

Dans son article de référence, Marvin Minsky les définit comme suit :

“A frame is a data-structure for representing a stereotyped situation [...]. Attached to each frame are several kinds of information. Some of this information is about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed. [...] Collections of related frames are linked together into frame-systems.” M. Minsky [Minsky, 1975]

Plus concrètement, les systèmes à base de schémas sont composés de structures de données (les schémas) combinant des informations déclaratives et procédurales avec des relations internes prédéfinies [Barr & Feigenbaum, 1982 (p.158)].

Dans leurs mises en œuvre, ils reposent sur quatre notions de base :

- > Les schémas : ils permettent de structurer des données complexes. Selon les approches, les schémas produisent un système uniforme (comme SRL) avec un seul type de schéma ou un système avec une dualité classe-instance (comme Unit).
- > Les attributs : associés aux schémas, ils permettent de définir la structure des données.
- > Les facettes : associées aux attributs, elles définissent leur sémantique.
- > Les relations : elles permettent d'associer des schémas et elles possèdent une sémantique d'héritage (les schémas héritent en général des attributs des autres schémas auxquels ils sont liés).

Remarque : certains auteurs trouvent des similitudes entre les systèmes à base de schémas et les langages de programmations orientés objets : tous deux représentent des classes (les schémas) avec des attributs. Néanmoins il existe de nombreuses différences (relations d'instanciation, mode d'activation des méthodes et des attachements procéduraux, etc.).

3.3.3 Présentation de SNDF

Pour rappel, notre problématique est de proposer un formalisme pour d'une part décrire les concepts du domaine et structurer l'espace selon ces concepts et d'autre part, pour définir des propriétés graphiques pour les différents éléments de l'espace (concepts, objets, relations, etc.).

Bien que notre objectif soit de décrire des connaissances, nous n'avons pas utilisé OWL et RDF¹. La première raison est qu'ils sont avant tout des formats d'échange. De plus, OWL ne permet pas de décrire des connaissances semi-formelles. RDF, quant à lui, ne permet pas de prendre en compte les contraintes spécifiques à notre problématique comme la définition et la manipulation de plusieurs théories ainsi que l'association d'actions aux entités de l'espace informationnel *structuré*. C'est donc pour ces raisons que nous avons choisi de proposer notre propre formalisme de description : SNDF pour Semantic Network Description Formalism. Néanmoins nous pouvons, à fin d'échange, traduire nos descriptions en RDF.

SNDF permet de décrire des espaces informationnels *structurés* par une variante des systèmes à schéma présentant deux facettes :

- > Les connaissances à cartographier :
 - > La modélisation : SNDF décrit un ensemble de concepts (entités particulières) et de relations entre ces concepts ;
 - > L'espace informationnel *structuré* : SNDF décrit un ensemble d'entités et leurs relations selon les concepts de la modélisation.
- > Les propriétés de représentations :
 - > Descriptions de représentations : SNDF permet de décrire pour chaque élément de l'espace un ensemble d'informations pour construire ses représentations (les structures visuelles).
 - > Descriptions des interactions : SNDF permet de décrire pour chaque élément de l'espace une liste d'actions et leur libellé pour décrire leurs interactions (via leurs représentations ; les structures visuelles).

Ces deux facettes de SNDF sont présentées dans les sections suivantes.

a. Décrire les connaissances du domaine

Les connaissances décrites selon notre formalisme ont pour finalité d'être représentées dans une carte. Nous n'avons alors pas besoin d'effectuer de traitement de type inférences sur ces connaissances. De plus, les formalismes dits « formels » n'introduisent pas de distinction entre prédicats et attributs or nous avons besoin de cette nuance pour exprimer la sémantique d'un domaine.

Par conséquent, nous pouvons écarter les formalismes « formels » (comme les ontologies ou les logiques) et orienter notre choix vers les théories de description semi-formelles. Ces dernières ont l'avantage de permettre de décrire tous types d'entités et de relations tout en proposant une approche souple et une description naturelle des choses.

Plus précisément, nous avons choisi une théorie basée sur les réseaux sémantiques. De plus, nous avons réutilisé la possibilité offerte par les systèmes à bases de schémas d'associer

¹ <http://www.w3.org/RDF>

des propriétés aux entités (correspondant aux « slots »). Mais contrairement aux systèmes à base de schémas, nous n'avons pas introduit ici la notion d'attachement procédural mais plutôt la notion de méthode (au sens des langages à objets) pour associer des actions aux entités manipulées.

Ainsi, les connaissances décrites avec SNDF sont exprimées par un ensemble d'entités et un ensemble de relations dont les tuples décrivent les entités en relations. Ensuite, nous avons étendu notre théorie pour permettre à notre formalisme de définir des ensembles d'entités afin de spécifier un domaine pour chaque relation (contraignant ainsi la nature des entités dans la définition des tuples).

Les connaissances manipulées avec notre formalisme sont décrites par un réseau sémantique, c'est-à-dire sous la forme d'un ensemble d'entités liées par des relations (arcs orientés étiquetés).

Pour SNDF, un réseau sémantique est défini par :

- > Un ensemble d'entités (possédant des attributs),
- > Un ensemble de relations entre les entités,
- > Un ensemble d'ensembles contenant tous les ensembles d'entités (utilisés pour définir le domaine des relations).

Le schéma suivant illustre l'utilisation de notre formalisme pour construire une modélisation puis pour structurer l'espace informationnel selon cette modélisation :

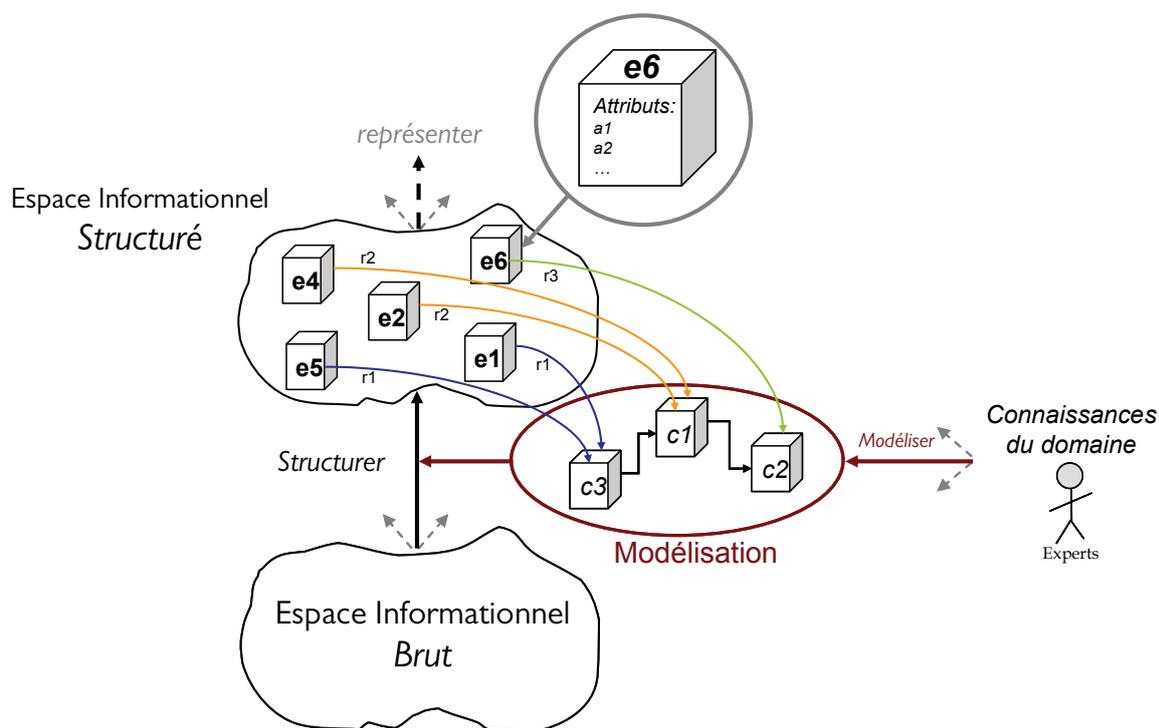


Figure 71 - SNDF pour décrire des modélisations et des espaces informationnels structurés.

Chaque description (modélisation ou espace structuré) est construite avec les notions d'entité (ou schéma), d'attribut, d'ensemble et de relation. Ces notions sont décrites dans les sections suivantes.

i. Entité et attribut

Pour les réseaux sémantiques, les entités peuvent représenter différents types de connaissance : un concept, une idée, un objet, une action, un mot, un processus, une situation, etc.

En revanche dans notre formalisme, les entités ne sont pas que des simples nœuds. Il est possible de caractériser chaque entité par un ensemble d'attributs possédant des valeurs. Les entités peuvent alors être qualifiées de « schémas ». Dans ce cas, le réseau devient un réseau de schémas. De plus, les attributs sont hérités par les relations prédéfinies « est un » et « est instance de ».

Remarque : les attributs sont parfois qualifiés de « relations internes » au schéma pour les distinguer des « relations externes » entre schémas.

ii. Ensemble

La définition de relations, comme décrit dans la section suivante, nécessite de pouvoir définir si besoin est ses propres ensembles d'entités sur lesquels reposent la définition des relations. C'est pourquoi, notre formalisme permet de définir des ensembles et de leur ajouter des éléments (entité, ensemble ou relation).

À la création de tout nouveau réseau sémantique, notre formalisme prédéfini trois ensembles particuliers :

- > « E » : il contient l'ensemble des entités. À la création de toute nouvelle entité, elle est nécessairement ajoutée à cet ensemble et ne peut en être enlevée (sauf si l'entité est supprimée). Une entité appartient donc à au moins un ensemble.
- > « R » : cet ensemble contient toutes les relations du réseau. Toute nouvelle relation est nécessairement ajoutée à cet ensemble. Une relation ne peut appartenir à aucun autre ensemble.
- > « S » : cet ensemble contient tous les ensembles du réseau (l'ensemble des ensembles). Tout nouvel ensemble créé est nécessairement ajouté à cet ensemble et ne peut en être enlevé (sauf si l'ensemble est supprimé). A l'exception de S, les ensembles ne peuvent contenir que des entités ou des relations.

Il existe différents types d'entités selon la nature de la connaissance qu'elles représentent : concept, objet (au sens d'instance d'un concept) et indéterminé (ni concept, ni instance). Toute entité d'un modèle est nécessairement d'un de ces trois types. L'ensemble E des entités est donc partitionné (au sens mathématique du terme) en trois ensembles : « Concepts », « Things » et « Unknown ». Par défaut une entité est de type indéterminé.

Les ensembles prédéfinis (en plus de E, R et S) sont donc :

- > « Concepts » : regroupe les entités déclarées comme étant des concepts ;
- > « Things » : ensemble des entités déclarées comme étant des instances.

iii. Relation

Les entités sont liées entre elles par des relations. La richesse d'un formalisme de description de connaissances dépend directement de la diversité des relations proposées. C'est pourquoi, notre formalisme permet de définir de nouvelles relations et il possède deux relations prédéfinies :

- > « est un » : liant un concept, c'est-à-dire une entité de type concept, à un autre concept ;
- > « est instance de » : liant une « instance », c'est-à-dire une entité de type « things », à un concept ;

Notre formalisme permet également de définir d'autres relations. Chaque relation possède un domaine et un ensemble de tuples. Le domaine permet de contraindre la nature des entités mise en relations. Il est défini par un produit cartésien d'ensembles. Les tuples sont définis par un vecteur d'entité, où chaque entité appartient à l'ensemble correspondant dans le domaine de la relation.

Par exemple avec la relation binaire « est un », son domaine est défini par le produit cartésien de l'ensemble des concepts par lui-même. Ainsi, la relation « est un » ne peut lier entre eux que deux concepts.

La définition de relations nécessite donc de pouvoir définir (si besoin est) de nouveaux ensembles d'entités.

Remarque : notre formalisme est dédié à des descriptions semi-formelles de connaissances. Ce qui signifie qu'il est faiblement contraint. Ainsi, il vérifie que la définition des relations est bien respectée lors de la mise en relation d'entités. Cependant, si la relation « est un » est bien définie sur le produit cartésien Concepts x Concepts et la relation « est instance de » sur Things x Concepts, le formalisme ne contraint pas leur utilisation. Ainsi, même si « l'instanciation multiple dans une même théorie » est à éviter, un objet (entité de l'ensemble Things) peut-être « est instance de » plusieurs concepts. Ce cas ne doit pas être confondu avec celui où un même objet peut être appréhendé par rapport à des théories différentes (modélisations) et peut-être « est instance de » de plusieurs concepts pris dans des modélisations différentes. De même, un concept peut être « est un » de plusieurs autres concepts, même si cela est à éviter d'un point de vue conceptuel.

iv. Partage d'instances entre modélisations

Notre processus ainsi que notre formalisme permettent de manipuler plusieurs modélisations simultanément. Chaque modélisation permet d'exprimer un « point de vue théorique » sur un domaine. Ainsi un même « objet » (instance), s'il est lié à un seul concept par la relation d'instanciation dans une modélisation donnée (ou du moins devrait l'être), peut être lié à un autre concept par la même relation dans une autre modélisation. Le même objet peut être appréhendé (conceptualisé) selon plusieurs points de vue théoriques différents. Notre théorie permet donc de partager les objets entre les modélisations.

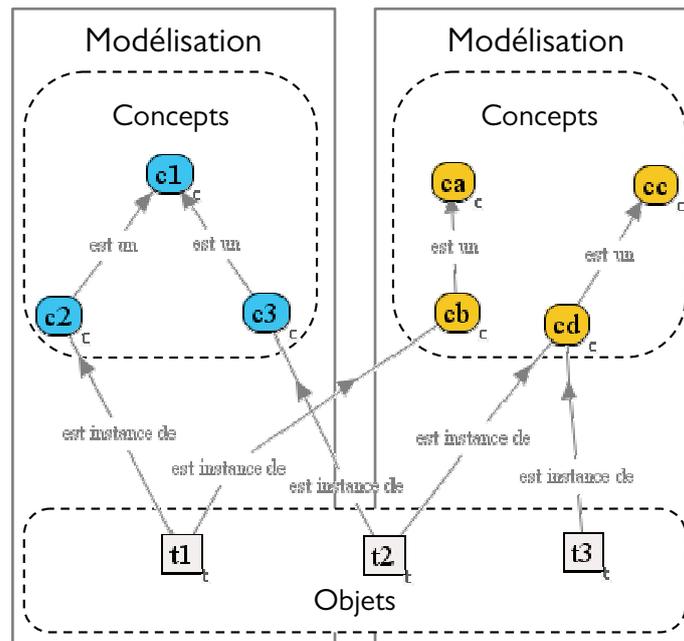


Figure 72 - SNDF : notre formalisme permet le partage d'objets entre deux modélisations.

b. Décrire les propriétés de représentations

Une fois structuré par la modélisation, l'espace doit être représenté puis visualisé pour produire des cartes. Les éléments de l'espace sont alors visualisés dans différentes cartes et l'utilisateur peut interagir avec leur représentation.

La construction de l'espace de représentation permet d'associer des structures visuelles aux éléments de l'espace *structuré* (les descriptions d'espaces informationnels *structurés* vont servir à la construction d'espaces informationnels *représentés* puis *visualisés*). Pour assurer une cohérence de représentation et d'interaction entre les cartes, la cartographie attribue généralement aux éléments de l'espace des propriétés communes à leurs structures visuelles dans toutes les cartes.

Cette facette de notre formalisme permet d'attribuer une représentation intrinsèque aux éléments de l'espace informationnel.

SNDF permet de décrire ces propriétés pour chaque élément de la modélisation et de l'espace informationnel *structuré*.

i. Les descriptions de représentations

Lors d'une opération de représentation, des structures visuelles sont construites par une opération pour représenter des éléments de l'espace informationnel. Chaque structure visuelle peut être décrite par un ensemble de valeurs associées aux différentes variables visuelles (comme une forme, une taille, une couleur, etc.). Ces couples (variable et valeur) doivent être précisés lors des opérations qui permettent de construire les structures visuelles.

Par exemple, imaginons que l'on souhaite construire une carte contenant un ensemble d'entités représentées chacune par un cercle rouge. Dans ce cas, on utilise une opération de représentation pour construire les structures visuelles de chaque entité en lui passant en paramètre la couleur et la forme (cercle rouge). Ce mécanisme est simple et la description est courte.

Nous savons et c'est souvent le cas, qu'il est possible de construire plusieurs cartes différentes d'un même espace informationnel *structuré*. Pour assurer une certaine cohérence de représentation entre les cartes, une entité possède généralement une représentation similaire (mais pas nécessairement identique) entre les différentes cartes.

Par exemple imaginons la cartographie décrite par le schéma suivant :

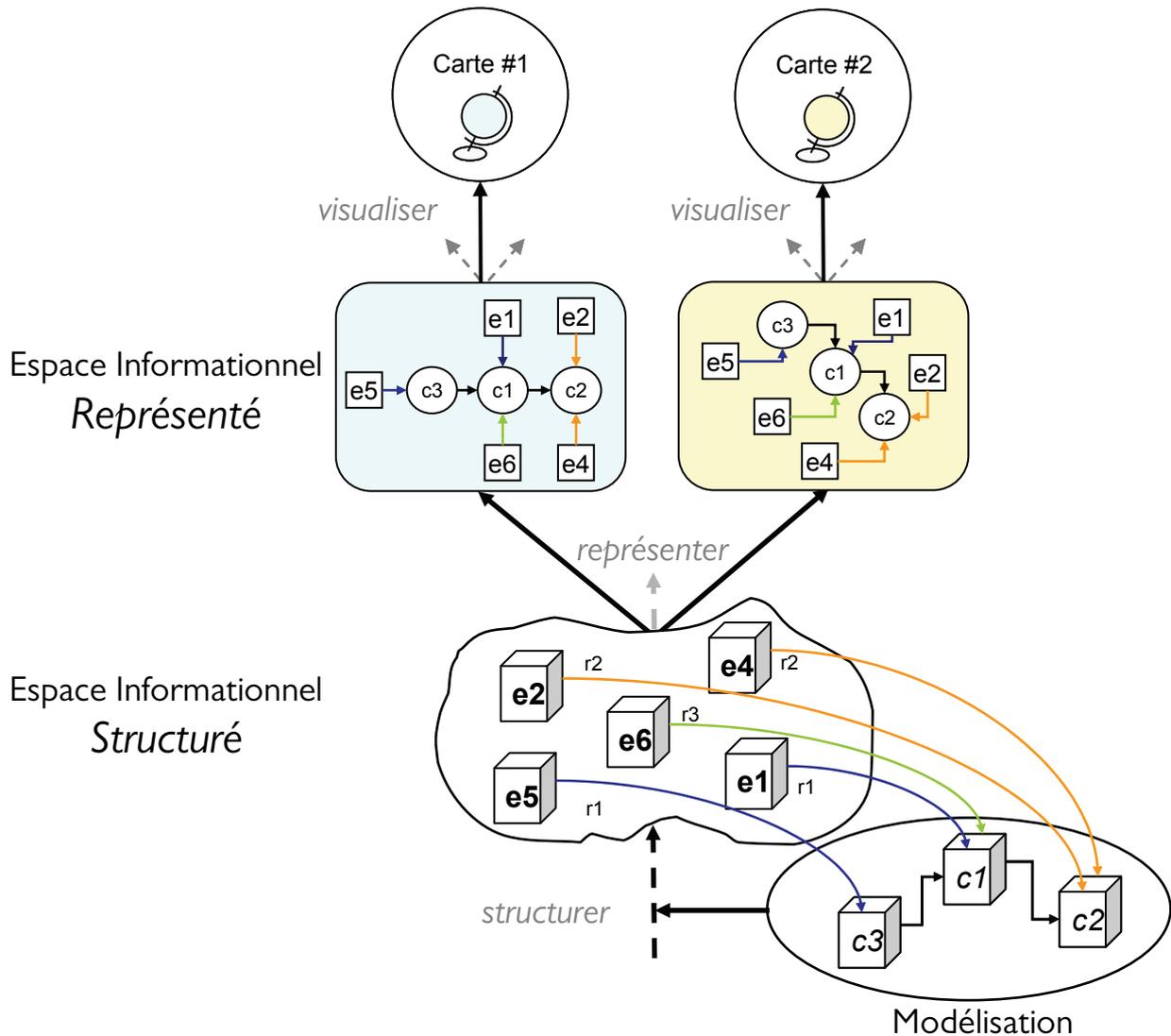


Figure 73 - Les cartes ont des structures visuelles similaires.

Dans cet exemple, on souhaite produire deux cartes différentes du même espace informationnel. Ces deux cartes possèdent chacune un espace informationnel *représenté* différent obtenu par une série d'opérations de représentation. Néanmoins, les entités ont des représentations similaires dans toutes les cartes. Par exemple, l'entité « e1 » est représentée par un carré noir dans les deux cartes et les tuples de la relation « r2 » sont représentés en orange dans les deux cartes.

Notre objectif est alors de garantir cette cohérence graphique entre les cartes pour ne pas perturber les utilisateurs. De plus, pour construire les structures visuelles de chaque carte, il est nécessaire d'effectuer une série d'opérations de représentation paramétrée par la

description des structures visuelles souhaitée (forme, taille, couleur). Si l'on construit plusieurs cartes d'un même espace, pour chaque carte et pour chaque élément à représenter, il faut utiliser des opérations de représentation paramétrées avec les mêmes valeurs.

Pour parvenir à notre objectif – garantir une cohérence graphique entre les différentes cartes et faciliter les opérations de cartographie –, SNDF permet d'associer aux éléments du réseau sémantique une liste de descriptions de représentation. Chaque description de représentation correspond alors à un ensemble de valeurs pour les variables visuelles.

Elles ne sont pas les représentations mais uniquement les différentes valeurs de quelques variables visuelles qui serviront pour construire chaque représentation (les structures visuelles). Le nombre de descriptions n'est pas limité étant donné qu'un élément (entité, relation ou tuple) peut avoir plusieurs représentations (à raison d'une ou plusieurs représentations par carte dans une ou plusieurs cartes).

Lors des opérations de création des structures visuelles, il suffit alors de préciser la description à utiliser.

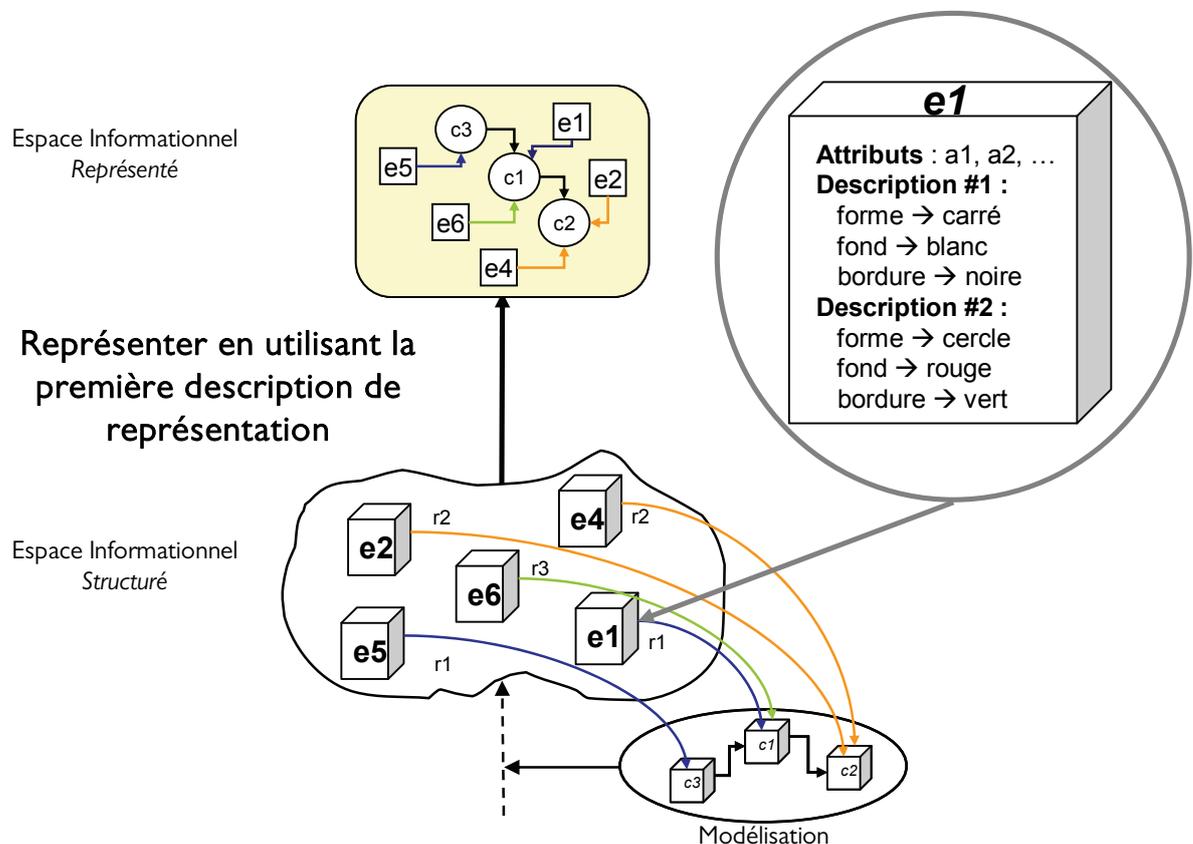


Figure 74 - Les éléments possèdent une liste de descriptions de représentations.

Notre formalisme offre donc la possibilité d'associer une liste de descriptions de représentations à tous les éléments du réseau sémantique (entité, relation, tuple, etc.), et ainsi de garantir une cohérence graphique entre les différentes représentations d'un même élément de l'espace informationnel.

ii. Les interactions par actions

Dans la section précédente, nous avons vu que pour des raisons de cohérence, un élément de l'espace informationnel possède généralement une représentation similaire dans toutes les cartes. Cette constatation est aussi valable pour les actions qui leur sont associées.

Avec l'informatique, les cartes sont devenues interactives et l'utilisateur peut interagir avec ses différents éléments (ces interactions se font généralement avec la souris). C'est par exemple le cas, lorsque l'on souhaite ouvrir une ressource externe associée à une entité (comme un document ou une page Internet). À chaque interaction correspond une action et par conséquent à chaque élément est associée une liste d'actions.

Ce mécanisme d'interaction est extrêmement fréquent et il est surtout très intuitif. On le retrouve par exemple dans les explorateurs de fichiers : le système de fichier est représenté dans un panneau latéral par une liste indentée et le contenu du répertoire courant est représenté dans un panneau central par des icônes. Chaque panneau correspond à une carte et les actions accessibles par l'utilisateur pour un objet donné (répertoire, fichier, etc.) sont communes aux deux cartes (les menus contextuels sont identiques).

Ces interactions sont donc assimilables à une liste d'actions proposées à l'utilisateur et cette liste est partagée par toutes les représentations d'un même élément. Par conséquent, SNDF permet de définir pour chaque élément du réseau sémantique une liste d'actions. À chaque action est associé un libellé descriptif.

3.3.4 Synthèse

Nous avons vu SNDF (Semantic Network Description Formalism) notre formalisme de description pour décrire d'une part, les connaissances du domaine en un ensemble d'entités et de relations entre ces entités et d'autre part des propriétés de représentation pour ces entités et ces relations.

Pour décrire les connaissances du domaine, SNDF a un double objectif. D'une part, il permet d'explicitier la sémantique du domaine en décrivant une modélisation et d'autre part, il décrit l'espace informationnel *structuré* selon cette modélisation. SNDF repose sur les réseaux sémantiques et les systèmes à base de schémas. Des premiers, il exploite la capacité de décrire des connaissances par un ensemble d'entités et de relations entre ces entités et des seconds, il reprend la possibilité d'ajouter aux entités un ensemble d'attributs valués. Pour véhiculer au mieux la sémantique du domaine dans la carte, notre formalisme permet aussi de définir un domaine (produit cartésien d'ensembles) pour chaque relation.

Pour décrire les propriétés de représentation pour les éléments du réseau sémantique, SNDF associe aux éléments (entités, relation ou tuple) des informations dédiées à la cartographie comme une liste de descriptions de représentation (la description des structures visuelles) et une liste d'actions pour les interactions.

Pour synthétiser, notre formalisme permet de décrire l'espace informationnel *structuré* par une variante de réseau sémantique présentant deux facettes :

- > Les connaissances à cartographier :
 - > La modélisation : SNDF décrit un ensemble de concepts (entités particulières) et de relations entre ces concepts ce qui permet de gérer différentes théories pour modéliser un même ensemble d'objets ;

- > L'espace informationnel *structuré* : SNDF décrit un ensemble d'entités et leurs relations selon les concepts de la modélisation.
- > Les propriétés de représentations :
 - > Descriptions de représentations : SNDF permet de décrire pour chaque élément de l'espace un ensemble d'informations pour construire ses représentations (les structures visuelles).
 - > Descriptions des interactions : SNDF permet de décrire pour chaque élément de l'espace une liste d'actions et leur libellé pour décrire leurs interactions (via leurs représentations ; les structures visuelles).

SNDF est notre formalisme de description dédié à la cartographie sémantique.

3.4 La cartographie sémantique pour l'exploration

Nous avons vu dans la section précédente comment décrire grâce à notre formalisme une modélisation et un espace informationnel *structuré*. L'objectif de cette section est donc de présenter notre démarche pour cartographier de tels espaces. Plus particulièrement, nous nous intéresserons à la cartographie sémantique pour l'exploration d'une base de connaissances annotée sémantiquement par une ontologie de domaine.

3.4.1 Positionnement de la problématique

Une base de connaissances annotée sémantiquement par une ontologie de domaine constitue un espace informationnel *structuré* par la sémantique du domaine. L'ontologie définit un ensemble de concepts ainsi que leurs relations. Dans notre cas, il existe une relation unique entre les concepts, celle de généralisation/spécialisation. L'ontologie explicite et décrit la sémantique du domaine. Pour se replacer par rapport à notre processus, l'ontologie constitue la modélisation du domaine et l'espace informationnel est structuré par cette ontologie.

La problématique est donc de définir la cartographie sémantique pour l'exploration d'un espace informationnel *structuré* selon une ontologie de domaine.

Si l'on reprend notre processus, on peut situer le problème comme suit :

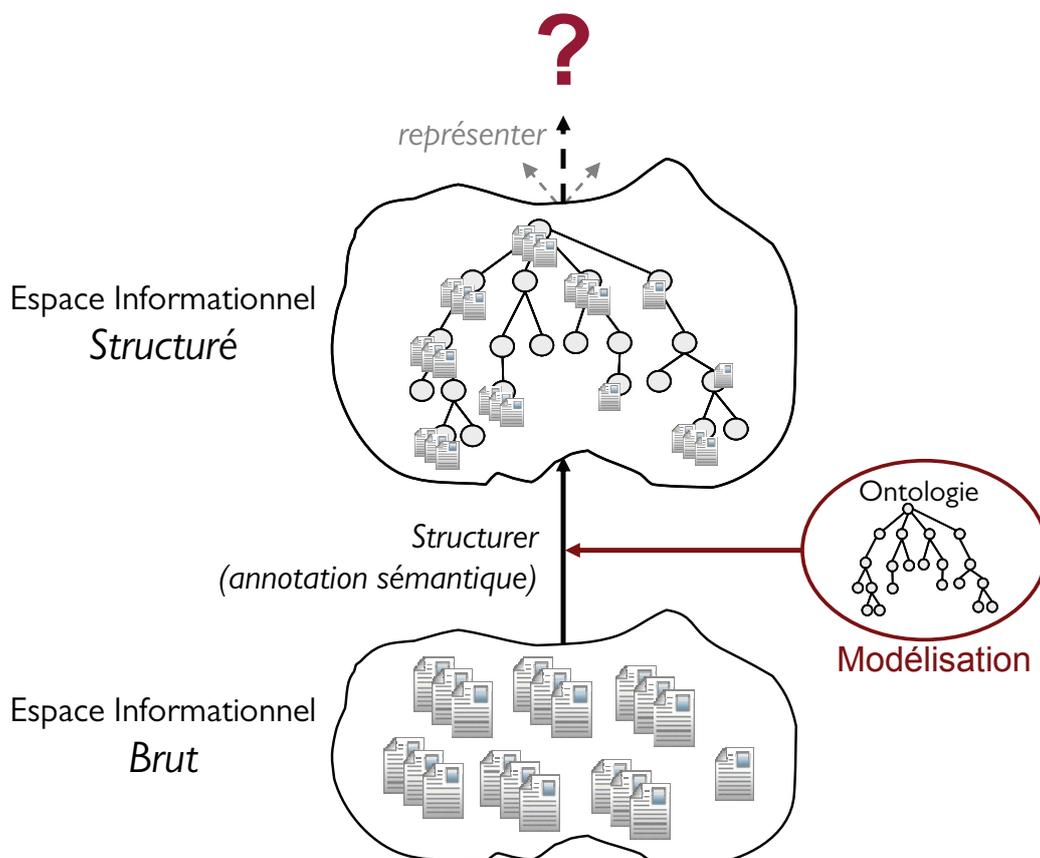


Figure 75 - Positionnement de la problématique par rapport à notre processus.

La problématique est de définir les paradigmes de représentation et de visualisation pour répondre à nos besoins.

Pour la résoudre, nous avons appliqué une démarche en deux étapes. La première étape a été d'étudier et d'évaluer les différents paradigmes de cartographie existants. Pour cela, nous avons au préalable défini avec l'aide des utilisateurs les fonctionnalités attendues. La seconde étape a consisté quant à elle à réaliser plusieurs cartes interactives dont l'analyse des retours d'expérience nous a permis d'identifier de nouveaux critères pour définir une cartographie sémantique dédiée à l'exploration.

Le résultat a été la spécification et la réalisation de deux nouveaux paradigmes de visualisation dédiés à la cartographie de bases de connaissances annotées sémantiquement : le « EyeTree » et le « RadialTree ».

3.4.2 Contexte et besoins

Cette étude a été menée sur différentes bases de connaissances constituées de fonds documentaires indexés sur une ontologie de domaine. L'utilisation d'une ontologie permet d'indexer et de classer les éléments de la base de connaissances. L'indexation repose sur l'analyse des contenus textuels (et péris textes ou métadonnées dans le cas des documents multimédias) au regard du vocabulaire associé à l'ontologie. Pour plus de détails sur les ontologies de domaine considérées ainsi que sur l'annotation sémantique voir la section sur « La démarche Ousia » à la page 160.

La structuration de l'espace informationnel sur l'ontologie permet de considérer chaque concept comme un répertoire virtuel contenant les éléments de la base de connaissances portant sur le concept. Dans le cadre de ces travaux, l'ontologie est une arborescence de concepts (hiérarchie simple) construite selon la relation de généralisation/spécialisation (voir à la page 160).

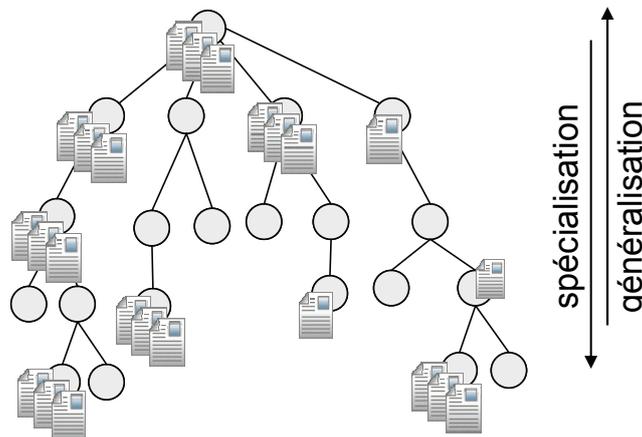


Figure 76 - Indexation des documents sur une ontologie de domaine.

Pour l'exploration de bases de connaissances, nous avons identifié plusieurs besoins que la cartographie doit satisfaire pour répondre à notre problématique :

Naviguer selon la sémantique du domaine : c'est-à-dire pouvoir accéder à l'information en parcourant les liens hiérarchiques entre concepts. Ainsi, l'exploration s'effectue selon la modélisation du domaine ce qui permet à l'utilisateur de la comprendre, de l'assimiler et de

l'exploiter et par voie de conséquence, il peut appréhender plus facilement l'ensemble de l'espace informationnel.

Proposer une vision à plusieurs échelles : permettre à l'utilisateur de s'approprier l'information nécessite de lui fournir les moyens pour l'appréhender dans sa globalité mais aussi dans ses particularités. C'est pourquoi, il est nécessaire d'offrir aux collaborateurs simultanément une vision globale et synthétique de l'espace informationnel et une vision particulière et détaillée de ce même espace.

Ces besoins sont l'expression de la difficulté qu'ont les utilisateurs à évoluer dans un espace informationnel important. Notre approche consiste donc à appliquer notre démarche de cartographie sémantique afin d'apporter des solutions à ce type de problèmes.

3.4.3 Cartographie de concepts et critères de cartographie

Le domaine de la visualisation de connaissances offre plusieurs paradigmes pour cartographier une structure hiérarchique de concepts. Ils sont généralement classables en deux familles : les cartes conceptuelles (concept mapping) et les cartes heuristiques (mind mapping) [Neumann *et al.*, 2005]. Ces deux familles sont décrites dans la section sur « Les approches par modélisations semi-formelles » page 174.

Ces techniques sont avant tout utilisées pour visualiser une structure conceptuelle et non pour aider les utilisateurs à explorer une base de connaissances. Particulièrement appréciées dans les scénarii d'apprentissage [Neumann *et al.*, 2005], ces cartes représentent les graphes sous forme de diagramme nœud-lien dans une vue uniforme. Le nombre de concepts simultanément affichables est par conséquent restreint.

a. Besoins et critères d'évaluation

Pour résoudre cette problématique, nous avons tout d'abord étudié et évalué les différents paradigmes de visualisation existants. Pour cela, nous avons au préalable défini avec l'aide des utilisateurs les fonctionnalités attendues. Notre approche repose sur la prise en compte des retours d'expérience (voir à la page 97). Cette approche est très différente des expérimentations directes pour évaluer la pertinence d'une visualisation comme avec [Kobsa, 2004; Pirolli *et al.*, 2003] ou [Barlow & Padraic, 2001].

De l'expression des besoins nous avons pu identifier, pour notre problématique, trois critères d'évaluation des différents paradigmes de visualisation à base d'arborescence de concepts.

1. **Visualisation de l'organisation des concepts** : dans la mesure où la conceptualisation du domaine joue un rôle central dans l'accès aux connaissances, il est important de pouvoir visualiser la structure globale de l'ontologie et ce quelle que soit sa taille. Étant donné que nous privilégions la relation hiérarchique de « généralisation – spécialisation », il est important que la disposition des concepts dans la carte respecte le mieux possible cette sémantique et ce dans un espace qui peut être réduit. La métaphore graphique à utiliser doit donc exprimer au mieux cette sémantique.
2. **Association d'informations aux concepts** : à chaque concept sont associées une liste de documents et une liste de termes. Il est donc nécessaire de pouvoir accéder et visualiser ces informations. La représentation des nœuds, en termes de variables graphiques comme la taille, la forme ou la couleur, doit être porteuse de sens. Un

utilisateur doit pouvoir accéder rapidement et intuitivement aux informations associées à un concept.

3. **Interaction & navigation** : l'utilisateur doit pouvoir naviguer au sein de son espace informationnel sans se perdre. À tout moment il doit pouvoir se localiser et identifier où il doit aller.

Nous avons ensuite mis en œuvre les principaux paradigmes connus (par exemple les arbres hyperboliques, les « Treemaps », etc.) pour les soumettre aux utilisateurs afin d'identifier les caractéristiques essentielles à prendre en compte.

b. Cartographies de concepts

Pour les domaines considérés lors de cette étude, il nous a été demandé de réaliser différents navigateurs graphiques d'accès aux documents en s'appuyant sur la modélisation du domaine (c'est-à-dire de pouvoir parcourir l'ensemble de la base en suivant les liens hiérarchiques de généralisation/spécialisation entre concepts). Un concept peut ainsi être interprété comme un « répertoire » contenant les documents qui s'y réfèrent.

Étant donné que la construction de la base de connaissances privilégie la relation de généralisation/spécialisation et en accord avec notre approche (voir à la page 94), nous avons retenu les techniques graphiques de type nœud-lien appliquées aux données hiérarchiques. En effet, ces techniques ont l'avantage de représenter explicitement la structure de l'arbre et par conséquent, elles expriment mieux la sémantique recherchée.

Cette contrainte nous a donc amené à écarter des techniques de cartographie par pavage (voir en page 72) comme les Tree-Maps les Cushion treemaps et les Beamtrees, ainsi que les disques d'informations. Les parties suivantes présentent les principaux paradigmes réalisés et étudiés pour cette étude.

i. *Prototype 1 : les listes indentées*

Principe. Une des visualisations d'arborescences les plus utilisées est l'affichage des répertoires telle qu'elle est utilisée pour les répertoires de fichiers dans les systèmes d'exploitation (Windows, Unix) ou dans des environnements de développements d'ontologies (Protégé¹, Oiled [Bechhofer et al., 2001]).

Cette visualisation exploite :

- > Une structure d'arbre dépliable pour représenter une hiérarchie de répertoires ;
- > Des icônes de dossier pour représenter les répertoires ;
- > Différentes icônes pour représenter les fichiers.

Généralement, pour la gestion de fichiers, la vue est découpée verticalement en deux avec à gauche la hiérarchie des répertoires et à droite, une zone pour afficher le contenu du répertoire.

¹ <http://protege.stanford.edu>

68 documents trouvés	
GRETH	
Condition de fonctionnement	
Fonctionnement du composant	
Fonctionnement du système	
Coefficient de performance	U ↓ ↕ 📄 p2003-10.pdf 23%
Concentration	U ↓ ↕ 📄 p2003-11.pdf 20%
Rendement thermique	U ↓ ↕ 📄 p2002-13.pdf 16%
Domaine d'application	
Bâtiment	
Air conditionné	U ↓ ↕ 📄 brevet-100701.pdf 8%
Chauffage	U ↓ ↕ 📄 p2002-8.pdf 7%
Climatisation	U ↓ ↕ 📄 p1999-3.pdf 6%
Eau chaude	U ↓ ↕ 📄 p1999-9.pdf 5%
Nouvelle technologie pour l'énergie	U ↓ ↕ 📄 p2001-13.pdf 4%
Industrie	U ↓ ↕ 📄 p2004-4.pdf 3%
Fluide	U ↓ ↕ 📄 brevet-230403.pdf 3%

Figure 77 - Prototype 1 : représentation en liste indentée (treeview simple).

Retour d'expérience. Cette technique est directement appropriable par l'utilisateur. Les répertoires sont étiquetés par les noms des concepts et le déploiement d'un nœud en nœuds plus spécifiques correspond bien à une interprétation naturelle de la relation de spécialisation. De plus, elle permet d'associer aux nœuds un nombre important d'informations qui peuvent être visualisées dans une zone dédiée (par exemple liste de documents).

Enfin, les interactions et la navigation au sein de l'arbre sont faciles et efficaces et l'utilisateur maîtrise son parcours qui reste visible à tout moment. Ceci est principalement dû au fait que les utilisateurs sont habitués à ce type de représentation.

Cependant dans le cadre d'applications concrètes où les ontologies peuvent être de taille importante, il devient difficile d'avoir une vue globale de la structure de l'arbre, *a fortiori* s'il est complètement déplié. L'utilisateur a alors du mal à naviguer au sein de la base de connaissances.

ii. Prototype 2 : les arbres de cônes

Principe. Afin de palier à la critique émise sur les listes indentées, nous avons réalisé un deuxième navigateur basé sur le paradigme des arbres de cônes (voir en page 68). Les arbres de cônes, tout comme les listes indentées sont des arbres de type nœud-lien.

Le principe consiste à dessiner l'ensemble de la hiérarchie en 3 dimensions (et non une vue partielle). Chaque nœud constitue le sommet d'un cône dont les fils se répartissent sur un cercle qui en constitue la base.

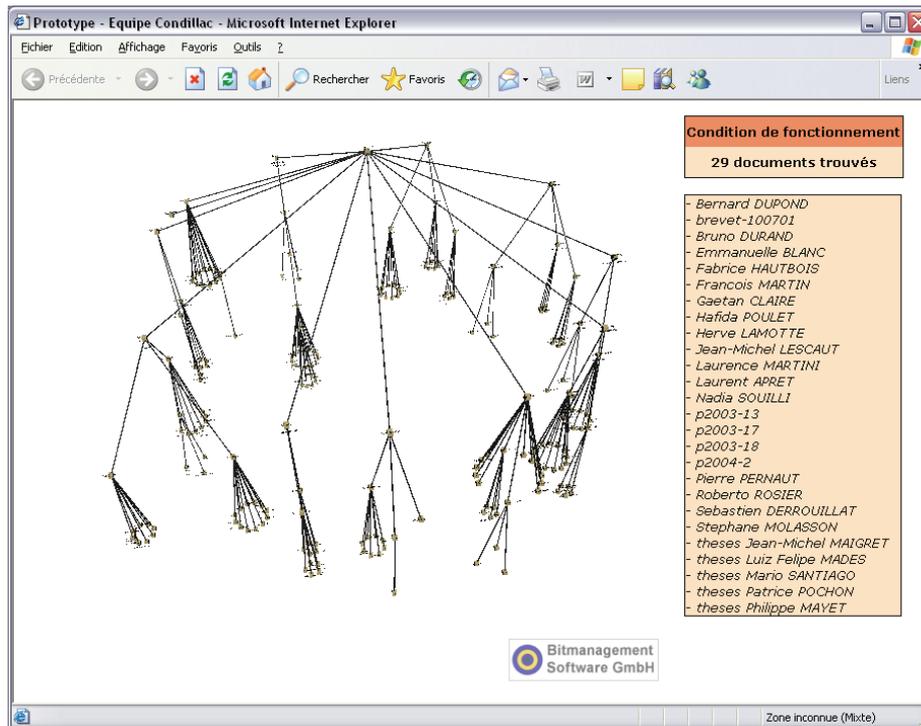


Figure 78 - Prototype 2 : les arbres de cônes.

Retour d'expérience. Si une telle visualisation donne un aperçu global de la structure de l'arborescence en termes de répartition des concepts et si elle semble séduisante par son interactivité, l'utilisateur est confronté à un phénomène d'occlusion et l'accès aux nœuds cachés par la structure nécessite de nombreuses manipulations de l'arbre.

Le parcours de la relation de généralisation/spécialisation est complexe. L'utilisateur n'est pas habitué à évoluer dans un espace informationnel à trois dimensions et se perd rapidement à l'intérieur d'un tel espace. L'effort cognitif est important et la prise en main de l'outil nécessite un long apprentissage.

iii. Prototype 3 : les arbres hyperboliques

Principe. L'idée ici n'est plus de vouloir visualiser de manière uniforme tous les nœuds, mais d'en visualiser certains de façon détaillée tout en permettant l'accès (visuel) aux autres nœuds.

Les arbres hyperboliques utilisent une technique graphique de visualisation non uniforme de type « fisheye » qui permet de placer dans la carte un nombre important de nœuds (voir à la page 69).

Ce type de carte utilise une géométrie non euclidienne : la géométrie hyperbolique. La représentation de la hiérarchie des concepts est un arbre radial visualisé sur un plan hyperbolique. Grâce à la géométrie de ce plan, l'utilisateur a l'impression que la taille des nœuds et la distance entre chaque nœud sont inversement proportionnelles à leur distance au centre du disque. Ainsi, les nœuds sont toujours visibles sinon accessibles, et il suffit à l'utilisateur de glisser au centre ceux qu'il souhaite voir plus en détails.

On obtient ainsi une vue de type « focus + context » où le focus est toujours au centre du disque.

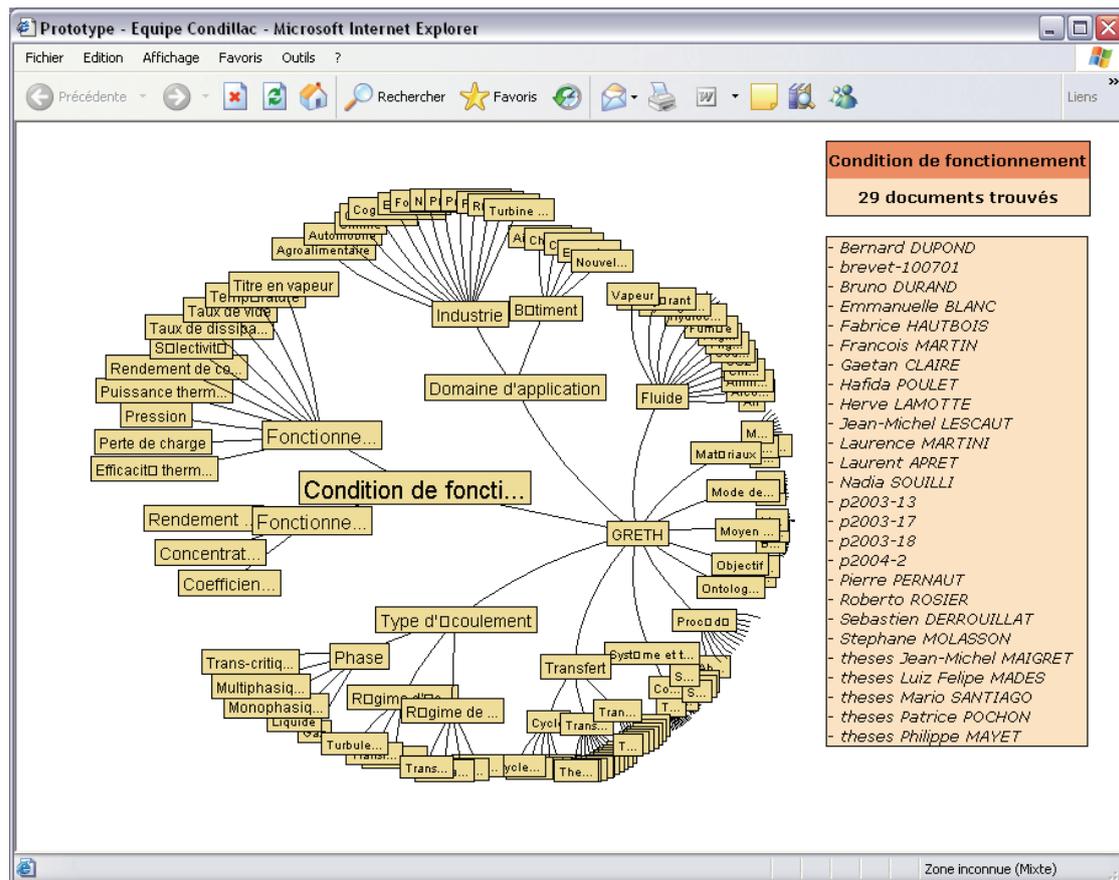


Figure 79 - Prototype 3 : les arbres hyperboliques.

Il existe des variantes en trois dimensions, mais elles ont l'inconvénient d'apporter des effets d'occlusion éliminant ainsi l'avantage de la « vision globale » offert par celle qui est en deux dimensions.

Retour d'expérience. Si de prime abord la forte interactivité des arbres hyperboliques séduit, elle souffre de plusieurs défauts qui peuvent en limiter sa réelle utilisation. Dû aux effets de la déformation, les étiquettes associées aux nœuds ne sont pas alignées et parfois se superposent. Mais c'est principalement son utilisation qui pose problème. En effet, lors de la manipulation de la structure, les éléments à la frontière de l'espace de visualisation se retrouvent projetés de façon « imprévisible ». Ces effets ont tendance à perturber l'utilisateur qui cherche en permanence à rétablir la situation engendrant un effort cognitif plus important et une prise en main assez délicate.

Ces effets - de projection sont dus à la géométrie utilisée. En effet, les éléments sont représentés dans un plan hyperbolique qui n'est pas commun à nos sens. C'est pourquoi, le résultat des transformations appliquées au plan n'est pas prévisible « naturellement ».

c. Critères de cartographie

Selon notre approche de la cartographie sémantique, nous avons analysé les retours d'expérience des utilisateurs. Ils nous ont permis d'identifier, dans le cadre de notre application, quatre critères principaux pour la réalisation d'une cartographie sémantique :

- > Utiliser une technique de type « focus + context » pour permettre à l'utilisateur de se concentrer sur certains éléments tout en facilitant l'accès aux autres éléments ;

- > Utiliser une géométrie euclidienne pour ne pas perturber la perception naturelle des manipulations du plan ;
- > Proposer une vue globale de la modélisation (l'ontologie) permettant à l'utilisateur de facilement appréhender l'ensemble de l'espace informationnel ;
- > Pouvoir parcourir la base de connaissances tout en gardant un point fixe de référence.

Suite à ces retours d'expérience nous avons également pu identifier deux profils distincts d'utilisateurs : les « novices » et les « experts ». Les « novices » ont une certaine connaissance des concepts du domaine sans connaître exactement leur organisation ; alors que les « experts » ont une bonne maîtrise de l'ontologie de leur domaine. Ces deux profils n'ont pas les mêmes attentes concernant la vision globale de l'ontologie. Pour les novices, il est nécessaire de leur proposer une carte permettant une appréhension globale de cette modélisation. En revanche, il est nécessaire de proposer aux experts une carte de l'ontologie pour se situer rapidement dans l'espace informationnel.

Ce constat nous a donc amenés à définir deux paradigmes de cartographies différents dédiés à la cartographie de bases de connaissances annotées sémantiquement : le « EyeTree » et le « RadialTree ».

3.4.4 EyeTree

Le « EyeTree » est notre premier paradigme de visualisation de concepts dédié aux « novices » du domaine. Il repose sur une représentation avec déformation de type « Fisheye Polaire » (voir à la page 80) avec une géométrie euclidienne offrant une vision globale de la structure conceptuelle.

Cette technique fait partie des techniques de représentation avec déformation [Leung & Apperley, 1994]. Pour cela, les nœuds sont répartis radialement dans l'espace euclidien (espace informationnel *représenté*) avant de subir une transformation via une fonction d'amplification continue appliquée aux coordonnées polaires des nœuds (espace informationnel *visualisé*).

La figure suivante illustre les opérations appliquées à l'espace informationnel *structuré* avant de le visualiser :

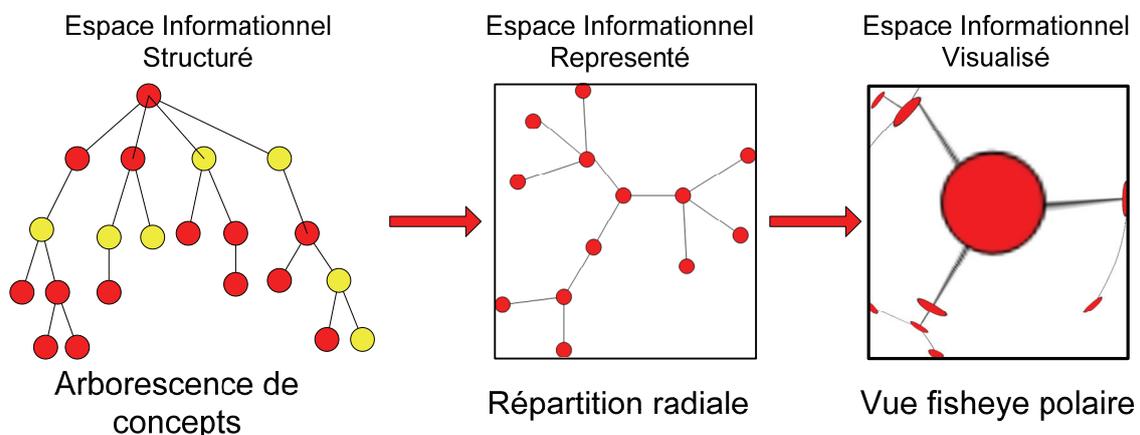


Figure 80 - EyeTree : opérations sur l'espace informationnel *structuré*.

Le résultat ressemble aux arbres hyperboliques, mais les interactions de l'utilisateur (par exemple les translations) sont appliquées à un plan euclidien. Elles sont donc « naturellement » prévisibles par l'utilisateur. La transformation étant linéaire, le résultat n'est pas perturbant pour les utilisateurs tout comme pour les « Perspective Wall ».

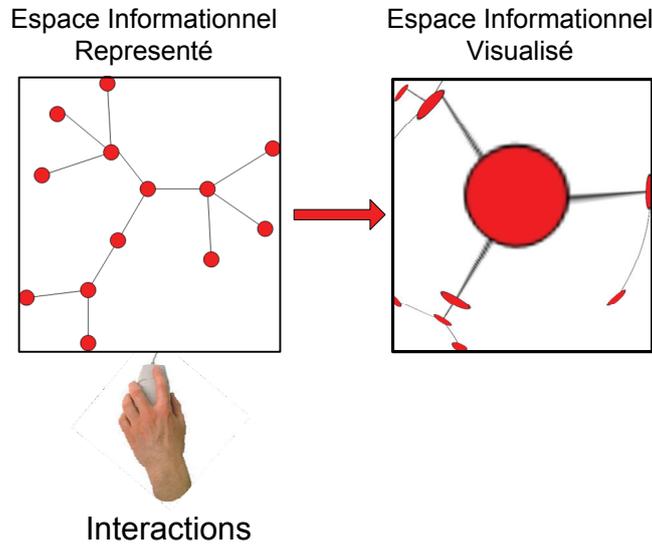


Figure 81 - EyeTree : interactions avec la carte.

Voici le résultat obtenu avec un nœud sélectionné et la liste des documents associés :

Figure 82 - EyeTree : exploration d'une structure de concepts.

La navigation avec l'EyeTree s'effectue en parcourant la relation de généralisation/spécialisation entre les concepts et en glissant la projection de l'ontologie (« tree ») sur la représentation sphérique (« eye »), d'où le terme de « EyeTree ». Une fois le concept sélectionné, la liste des informations associées (documents, données complémentaires, etc.) est affichée dans un volet latéral. Cette technique permet d'afficher la structure arborescente avec une déformation uniforme sur l'ensemble de l'arbre. La position relative des nœuds est toujours respectée et l'utilisateur perçoit uniformément l'ensemble de la structure. La carte obtenue possède des ressemblances avec le système Bookmap [Hascoët, 2001].

L'EyeTree permet aussi d'avoir une vue globale de la structure en faisant varier la distance entre les nœuds à l'aide de deux curseurs.

Enfin, les utilisateurs, face aux problèmes que pose l'utilisation des arbres hyperboliques et en particulier face au fait que toute modification locale entraîne des perturbations globales, ont exprimé la possibilité de pouvoir parcourir un ensemble d'éléments au sein d'un espace avec déformation tout en conservant une référence par rapport à un point fixe, en particulier par rapport à la racine de l'ontologie. Pour y parvenir, nous avons introduit les rotations du plan euclidien avec pour centre la racine de la structure hiérarchique. Cette interaction a pour conséquence de faire défiler tous les éléments du même niveau par rapport à un point fixe. Ainsi pour la recherche, il est possible de parcourir avec une seule interaction l'ensemble des sous-éléments d'un élément donné (ces rotations sont associées aux événements de la souris correspondant aux actions de la molette).

3.4.5 RadialTree

Par opposition à l'EyeTree, le « RadialTree » est plus particulièrement dédié aux experts du domaine.

RadialTree reprend les principales caractéristiques de l'EyeTree comme une cartographie globale de l'ontologie et l'utilisation d'une géométrie euclidienne. Cependant, le RadialTree n'applique pas de déformation globale à l'espace informationnel représenté.

En revanche, il propose une zone dite de « focus » permettant de magnifier une sous-partie de l'arbre. Cette zone est définie par une portion du disque dont l'angle est paramétrable par l'utilisateur.

Voici un exemple d'ontologie visualisée avec le RadialTree :

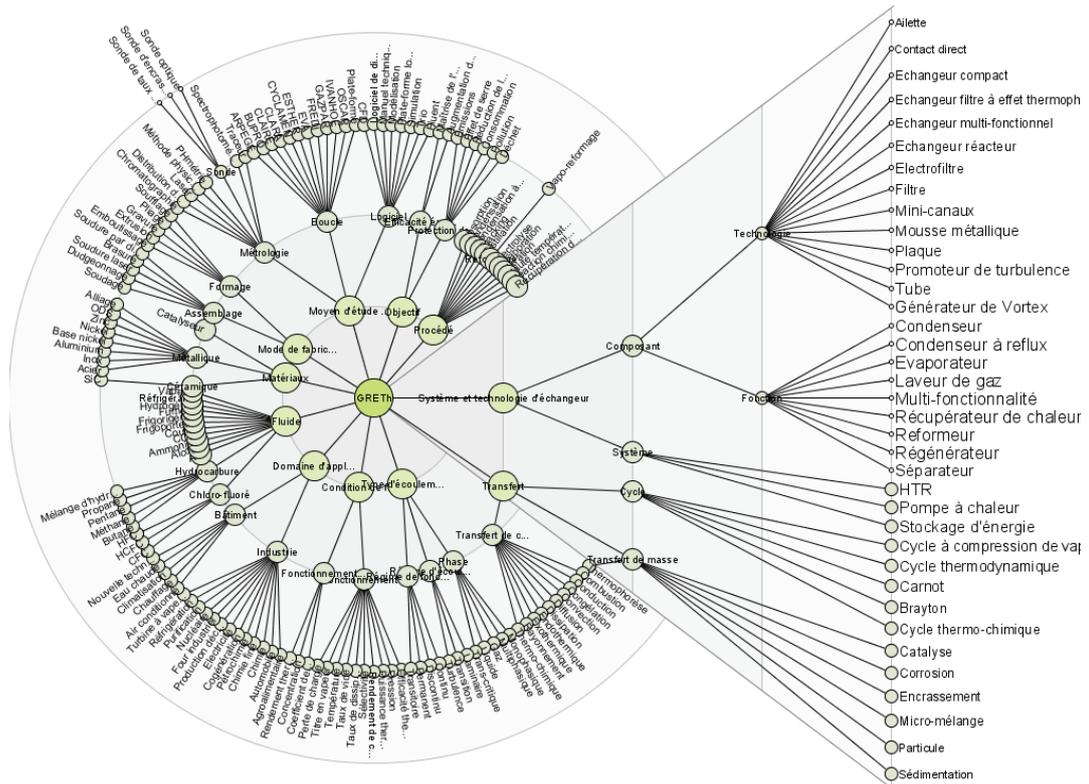


Figure 83 - RadialTree : exploration d'une structure de concepts.

Ainsi, les nœuds de l'arborescence sont répartis radialement dans l'espace euclidien. Les éléments situés dans la zone de focus sont magnifiés par une fonction d'amplification continue appliquée aux coordonnées des structures visuelles (les nœuds). De plus, les feuilles de l'arbre situées dans la zone de focus sont projetées sur l'extrémité droite de cette zone et leur taille subit une déformation pour donner un effet équivalent à celui des menus fisheye [Bederson, 2000] augmentant ainsi leur lisibilité. Il est à noter que la taille de la zone de focus ainsi que la profondeur de l'arbre à afficher sont réglables. Le principe de se focaliser sur un sous-arbre et de le magnifier sur la droite est proche de la navigation proposée par le système TreePlus [Lee et al., 2006].

Le paradigme RadialTree permet donc d'afficher une structure arborescente avec une déformation locale. La vue globale permet à l'utilisateur de se situer dans l'espace informationnel et la zone de focus permet de se concentrer plus efficacement sur un sous-arbre de concepts.

3.4.6 Synthèse

La problématique de cette partie était de définir la cartographie sémantique pour l'exploration d'un espace informationnel structuré selon une ontologie de domaine.

Pour résoudre cette problématique, nous avons appliqué une démarche en deux étapes. La première étape a été d'étudier et d'évaluer les différents paradigmes de visualisation existants. Pour cela, nous avons au préalable défini avec l'aide des utilisateurs les fonctionnalités attendues. La seconde étape a consisté quant à elle à réaliser plusieurs cartes interactives dont l'analyse des retours d'expérience nous a permis d'identifier de nouveaux critères pour définir une cartographie sémantique dédiée à l'exploration.

Le résultat a été la spécification et la réalisation de deux nouveaux paradigmes de visualisation dédiés à la cartographie de bases de connaissances annotées sémantiquement : le « EyeTree » et le « RadialTree ». La réalisation de ces deux propositions est présentée avec les autres réalisations associées à cette étude (voir à la page 189).

3.5 MDL : notre langage de cartographie sémantique

La mise en œuvre de notre méthode de cartographie sémantique a conduit à la proposition d'un processus dédié reposant sur une succession de changements d'espaces : *structuré* puis *représenté* et enfin *visualisé*. Notre formalisme permet de décrire des espaces informationnels *structurés* et nos paradigmes cartographient de tels espaces. Nous avons alors besoin d'un langage pour décrire une carte entendue comme la description de l'ensemble du processus de cartographie sémantique.

3.5.1 Nécessité d'un nouveau langage

Il existe de nombreux langages pour décrire des cartes. Il est possible de citer des langages comme Flash (langage propriétaire), SVG ou X3D (langages ouverts). Ces langages permettent de décrire des cartes comme un ensemble de structures visuelles (parfois ils autorisent à paramétrer les opérations de visualisation).

Par rapport à notre processus, ces langages ne permettent que de décrire des espaces informationnels *représentés*. Ils ne permettent pas de décrire l'espace informationnel et encore moins de décrire l'espace *structuré* par une sémantique de domaine. Par exemple, si l'on décrit une carte d'un ensemble de connaissances avec ces langages, la description ne peut contenir que des descriptions de structures visuelles représentant les connaissances (description d'un ensemble d'objets graphiques).

Le défaut des langages graphiques existants est qu'ils décrivent des cartes à un niveau d'abstraction très faible, au niveau graphique. Le lien entre la carte et les connaissances cartographiées est rompu. Les cartes perdent alors toute la sémantique associée à ces connaissances.

Pour illustrer l'importance de manipuler des connaissances (décrit par un espace informationnel *structuré*) prenons un organigramme dont voici un exemple de carte :

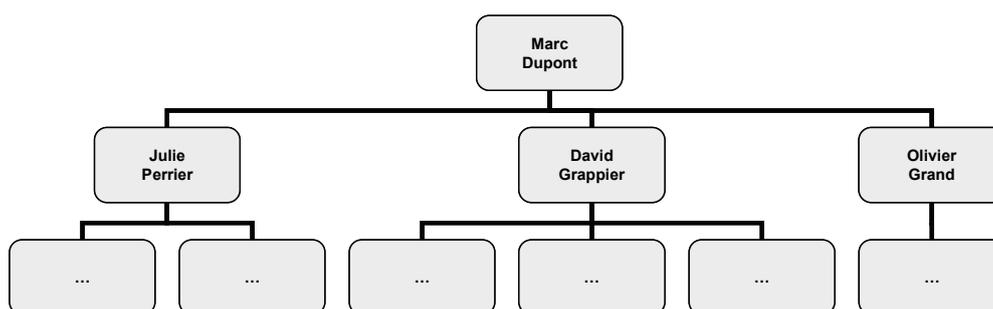


Figure 84 - Exemple d'une carte contenant un organigramme.

Si l'on souhaite décrire synthétiquement cette carte en langage naturel en se plaçant au même niveau d'abstraction que les langages graphiques existants, on obtient :

« La carte contient des rectangles pleins et arrondis. Ils sont de couleur noire et leur fond est gris clair. Ils sont reliés par des traits noirs. Sur chaque rectangle, il y a du texte ».

Il s'agit donc bien d'une description en langage naturel de l'espace informationnel *représenté* et plus particulièrement, des structures visuelles qu'il contient.

En revanche, si l'on souhaite décrire cette carte au niveau d'abstraction de l'espace informationnel *structuré*, on obtient la description (en langage naturel) suivante :

« La carte contient des entités correspondant à des personnes. Ces personnes sont reliées par une relation de responsabilité. Les personnes sont représentées par des rectangles arrondis gris clair avec le nom à l'intérieur et la relation de responsabilité est représentée par des arcs en noir. »

Cet exemple, volontairement très simpliste permet de mettre en évidence que le résultat final (la carte) est le même. Cependant dans le deuxième cas, la description est sémantiquement beaucoup plus riche. La carte contenant intrinsèquement les connaissances, il est alors possible pour l'utilisateur d'interagir (ou d'appliquer tout type d'opérations de cartographie) avec les représentations (déplacement d'objet, changements de point de vue, etc.) mais aussi avec les connaissances au travers de leur description par l'espace informationnel *structuré* (nouvelle représentation, classification, ...). Il est alors facile de comprendre l'intérêt de travailler sur des connaissances (représentant le monde réel) plutôt que de travailler sur leurs représentations (des objets graphiques).

MDL, notre proposition de langage de cartographie permet de décrire une carte en décrivant les connaissances à cartographier.

3.5.2 Principe de MDL

Pour répondre au besoin de cartographier sémantiquement des ensembles d'informations, nous avons créé notre propre langage. Nous l'avons nommé MDL pour « Map Description Language » ou « langage de description de cartes ».

L'objectif de MDL est de décrire une cartographie sémantique (entendue comme l'ensemble des opérations nécessaires à appliquer à un espace informationnel *structuré* pour créer une carte). D'un point de vue plus technique, MDL a été conçu pour être un langage à balises de la famille XML¹.

MDL est utilisé par le système Os Map, notre service Web de cartographie sémantique, décrit ultérieurement (voir en page 163). Os Map accepte des requêtes en MDL et il produit les cartes correspondantes.

MDL met directement en œuvre notre approche, notre processus ainsi que notre formalisme de description SNDF. En effet, selon notre processus, pour décrire une cartographie, il suffit de décrire l'espace informationnel à cartographier et les opérations à appliquer sur cet espace pour construire les cartes (il est possible de créer plusieurs cartes d'un même espace et de les utiliser simultanément).

Selon notre processus, la première étape débute par la modélisation du domaine pour expliciter sa sémantique. La seconde étape consiste à structurer l'espace informationnel selon cette modélisation.

¹ <http://www.w3.org/XML>

Expliciter la sémantique se fait en modélisant les connaissances du domaine qui sont à la fois explicites et tacites. La modélisation ne se limite donc pas à traiter l'espace informationnel (connaissances explicites). Il est nécessaire d'avoir recours aux experts du domaine (transfert des connaissances tacites). Il n'est donc pas possible de modéliser le domaine par un traitement automatique. Par conséquent, MDL permet de décrire des cartographies qui débutent par l'espace informationnel déjà structuré. La sémantique du domaine (la modélisation) est alors intrinsèque à la structure de l'espace (les concepts de la modélisation servent à décrire l'espace).

Toute cartographie décrite en MDL est composée de deux éléments : l'espace informationnel *structuré* (les connaissances) puis, la succession des opérations à leur appliquer pour construire les cartes. Bien que toutes les cartes ainsi construites partagent les mêmes connaissances à cartographier, leur construction est une cartographie à part entière. C'est pourquoi nous désignerons par « sous-cartographie » toutes les cartes d'un même espace informationnel *structuré* visualisées simultanément en une carte globale (comme avec les paradigmes de vues multiples).

Une cartographie MDL est donc composée de connaissances et de cartographies :

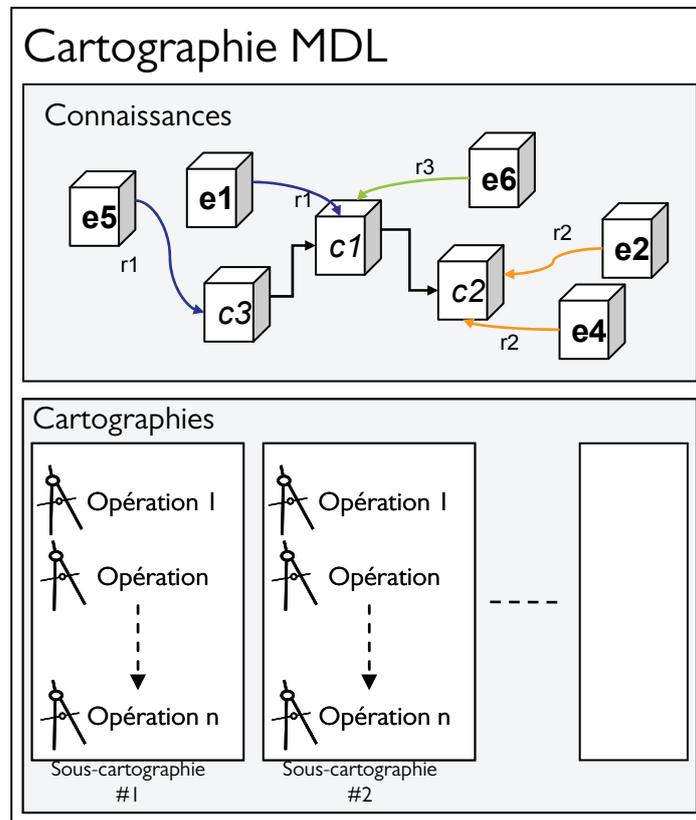


Figure 85 - Structure d'une cartographie décrite en MDL.

Le principe de MDL permet de ne pas perdre la sémantique associée aux connaissances cartographiées et par conséquent de travailler à un niveau d'abstraction plus élevé, celui des connaissances. L'utilisateur peut alors interagir avec les représentations (comme les langages existants) mais aussi avec l'espace informationnel *structuré*.

3.5.3 La structure du langage

Cette partie a pour objectif de décrire le principe et la structure des différentes parties du langage MDL. Pour plus de détails, la DTD complète est disponible en annexe. Cette DTD correspond à la description du langage MDL et plus précisément de ses éléments XML.

XML est un langage à balise qui offre de nombreux avantages mais pour des raisons de lisibilité, nous avons choisi de présenter MDL avec une syntaxe simplifiée de type BNF (expressions régulières). Ainsi, chaque égalité (« ::= ») correspond à la définition d'un élément du langage MDL avec le membre de gauche correspondant au nom de l'élément et le membre de droite correspondant à sa définition. Les virgules indiquent la séquence et les barres verticales « | » indiquent l'alternative. Les parenthèses permettent de créer des associations et les symboles « ? » (optionnel), « * » (plusieurs) et « + » (au moins un) indiquent les cardinalités.

Par exemple, pour les connaissances de la carte qui sont décrites par un ensemble de schémas et un ensemble de relations, on utilise la syntaxe suivante :

```
knowledge ::= (schema*,relation*)
```

Toute cartographie MDL est constituée des connaissances à cartographier puis de toutes les sous-cartographies (au moins une) correspondant à autant de cartes.

```
Map ::= (Knowledge, Cartographies)
```

La cartographie possède aussi un ensemble de valeurs permettant d'exprimer des métadonnées comme l'auteur, le titre, une description ou la durée de vie de la carte.

a. Les connaissances

Pour décrire les connaissances à cartographier nous utilisons SNDF, notre formalisme de description présenté précédemment (voir à la page 109). Ainsi, les connaissances sont exprimées par un ensemble de schémas (les entités à cartographier) et un ensemble de relations dont les tuples mettent les schémas en relations :

```
knowledge ::= (schema*,relation*)
```

i. Les schémas

Notre formalisme basé sur les systèmes à base de schémas permet d'associer des attributs aux entités à cartographier (correspondant aux « slots »). Ainsi, le cartographe (la personne qui va construire le fichier MDL dans notre cas) peut spécifier certaines particularités pour chaque entité.

Par défaut, certains attributs sont proposés comme un identifiant (entre autres pour simplifier la définition des tuples) et une étiquette.

Toujours selon notre formalisme, chaque schéma peut posséder des descriptions de représentation et une liste d'action que l'on nomme ici un « menu d'actions ». Ces

descriptions de représentation possèdent un identifiant (utilisées pour les passer en paramètre aux opérations lors de la cartographie) et différents attributs graphiques (comme la couleur).

Un schéma en MDL peut donc est décrit comme suit :

```
schema ::= id, attribut*,representationDescriptionSchema*,menu ?
representationDescriptionSchema ::= id, x, y, z, strokeSize, strokeColor,
backgroundColor, style, picture, font, ...
menu ::= (label,action)*
```

ii. Les relations

Une relation possède un nom et un ensemble de tuples. Chaque tuple est un n-uplets qui permet d'associer des schémas et il peut être décrit comme une liste ordonnée de noms de schéma. Pour simplifier le langage, en MDL les tuples sont définis par une liste d'identifiants de schéma.

Tout comme les schémas, les relations possèdent un ensemble d'attributs dont un nom et une description. De même, une relation peut posséder un menu et des descriptions de représentation qui vont servir pour la création des structures visuelles des différents tuples (dans la plupart des cas, les relations ne sont pas représentées graphiquement mais uniquement leur tuples).

```
relation ::= tuple*,attribut*,representationDescriptionRelation*, menu ?
tuple ::= idSchema*
representationDescriptionRelation ::= id,color,stroke, ...
```

Cette première partie présente comment décrire en MDL les connaissances à cartographier : un ensemble de schémas et un ensemble de relations. Il reste alors à voir la partie cartographie du langage qui permet de passer d'un espace informationnel *structuré* à une ou plusieurs cartes.

b. Les cartographies

Notre processus de cartographie permet de construire une ou plusieurs cartes d'un même espace informationnel : les points de vue simultanés. Chaque carte possède son propre processus de cartographie pour construire l'espace informationnel *représenté* puis l'espace informationnel *visualisé*.

Pour organiser toutes ces cartes, il existe plusieurs paradigmes. C'est par exemple le cas avec le paradigme « Overview + Details » où une carte présente l'ensemble des connaissances et une carte complémentaire présente en détails un sous-ensemble beaucoup plus restreint mais plus détaillé de connaissances.

MDL propose un mécanisme de description de mise en page pour organiser les différentes cartes d'un même espace informationnel inspirés des mécanismes de mise en page de composants du langage de programmation Java¹. Ce dernier offre un mécanisme descriptif relativement simple pour gérer une mise en page de différents composants.

En java, ces gestionnaires permettent de décrire dans un premier temps la disposition globale des différents composants dans l'interface (correspondant au type de gestionnaire) puis

¹ <http://java.sun.com/docs/books/tutorial/uiswing/layout/visual.html>

dans un deuxième temps, les gestionnaires se chargent d'ajouter effectivement les composants dans l'interface. La grande variété de type de gestionnaire permet de réaliser un nombre très important de dispositions différentes, largement suffisantes pour nos besoins.

MDL reprend ce principe de gestionnaire de mise en page et le schéma suivant en présente les principales dispositions :

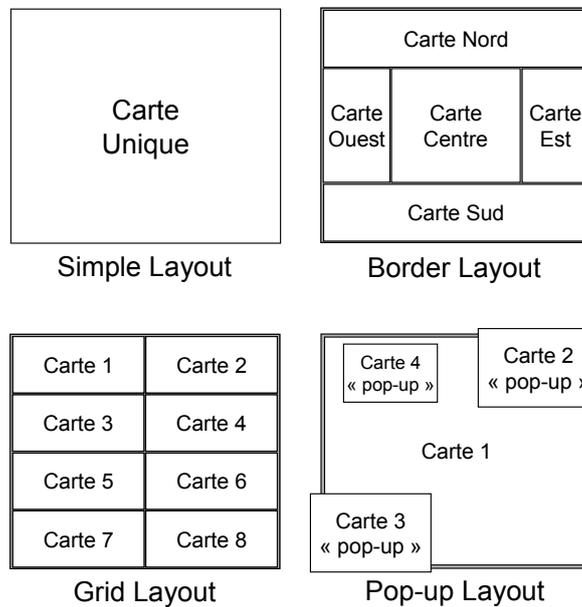


Figure 86 - Gestionnaires de mise en page de MDL.

En MDL, l'organisation des sous-cartographies est décrite comme suit :

```

Cartographies ::= cartography | borderLayout | gridLayout | popupLayout
borderLayout ::= center ?,north ?,south ?,east ?,west?
center: = north = south = east = west ::= cartography
gridLayout ::= row+
row ::= column+
column ::= cartography*
popupLayout ::= cartography*
    
```

i. La cartographie

Chaque (sous-)cartographie correspond à une vue sur l'espace informationnel. Cette partie décrit comment définir une (sous-)cartographie en MDL.

D'après notre processus, une cartographie peut être définie par un ensemble d'opérations à appliquer à l'espace informationnel *structuré* pour obtenir une carte. Chaque opération produit qu'un espace de niveau équivalent ou immédiatement supérieur ou inférieur. Il est donc possible de classer les opérations selon le niveau de l'espace en entrée et celui qui est en sortie.

Ainsi pour les besoins de MDL, nous avons défini plusieurs classes d'opération dont voici les principales :

Ordre	Classe d'opérations	Espace en entrée	Espace en sortie
1	Représentation	Structuré	Représenté
2	Transformations de Représentation	Représenté	Représenté
3	Visualisation	Représenté	Visualisé
4	Transformations de Visualisation	Visualisé	Visualisé

Tableau 6 - Classes d'opérations en MDL.

Pour garantir la cohérence de la cartographie, il est nécessaire d'avoir un certain nombre d'opérations qui garantissent de passer dans l'ordre par un espace *structuré*, un espace *représenté* et un espace *visualisé*. Si un espace de chaque famille (*structuré*, *représenté* et *visualisé*) seulement est utilisé, on peut alors parler de cartographie minimale. La cartographie minimale en MDL peut donc se limiter à deux opérations : une opération pour créer un espace *représenté* et une opération pour créer un espace *visualisé*, la carte.

La définition du langage MDL permet de contraindre la définition de toute cartographie dans l'ordre de ses opérations. On obtient alors la définition suivante :

```
Cartography ::=
  representationOperations,
  representationTransformationOperations?,
  visualisationOperations,
  visualisationTransformationOperations?
```

La définition de chaque classe d'opérations contient la définition de chaque opération. Les opérations possibles (indépendamment des classes) sont fonctions des opérations implémentées par le système qui souhaite exploiter MDL (MDL est un langage extensible). Cependant pour mieux comprendre MDL, nous pouvons prendre des exemples d'opérations issues de notre première mise en œuvre de MDL (dans Os Map décrit ultérieurement page 163) et plus particulièrement, des opérations qui permettent de créer des représentations. Ces opérations appartiennent à deux classes : la classe des opérations de représentations et la classe des opérations de transformations de représentations.

Dans la première version de MDL, nous avons deux opérations qui permettent de représenter un espace informationnel *structuré*.

La première que nous avons nommée « represent » permet de créer un espace *représenté* en prenant chaque élément de l'espace *structuré* (schémas et tuples) pour leur associer une structure visuelle. Cette opération prend un seul paramètre qui est la description de la représentation à utiliser (si elle existe) pour construire les structures visuelles.

La seconde opération, nommée « representTree » est similaire à la première opération, mais elle permet en plus de ne représenter que les schémas de l'espace *structuré* qui appartiennent à l'arbre dont la racine est constituée par un schéma (donné en paramètre) et toutes les relations de l'espace confondues. Cette dernière opération est utile par exemple lorsque l'on a un réseau de personnes liées par une relation de responsabilité. Cette opération permet alors de construire l'organigramme d'un collaborateur donné.

Pour la classe des opérations de représentations et pour les deux opérations considérées, on obtient la définition du langage MDL suivante :

```
representationOperations ::= (represent | representTree)
represent ::= idSchemaRepresentationDescription
representTree ::= idSchema, idSchemaRepresentationDescription
```

De même, nous avons défini quelques opérations permettant de modifier un espace *représenté* (création d'un nouvel espace *représenté* à partir d'un espace *représenté* donné). Nous ne présenterons ici que deux de ces opérations : « mapSchemaInStar » et « mapSchemaInTree ». Ces deux opérations ont pour objectif de répartir les représentations des schémas dans l'espace de la carte. La première opération permet de répartir les schémas radialement et la seconde en râteau. Les deux opérations admettent des paramètres comme le schéma (plus justement sa représentation) qui va constituer la racine de l'arbre, l'espace entre les structures visuelles (la longueur) et la profondeur maximale pour appliquer la répartition.

On obtient donc la définition suivante :

```
representationTransformationOperations ::= (mapSchemaInStar | mapSchemaInTree)*
mapSchemaInStar ::= idSchema, length, depth
mapSchemaInTree ::= idSchema, hSpace, vSpace
```

À noter que toutes les opérations de cette famille donnent un espace de la même famille. Par conséquent, il est possible d'appliquer un nombre illimité de fois successivement ces opérations sans altérer la cohérence globale du processus.

3.5.4 Exemples

En annexe (à la page 248), nous présentons deux exemples de carte décrites en MDL et cartographiées par Os Map, notre service web de cartographie sémantique. Os Map est décrit dans la partie sur nos réalisations.

Néanmoins pour illustrer notre langage dans cette section, nous présentons un extrait du premier exemple : « l'ontologie des êtres vivants ».

a. L'ordre de cartographie

```
<map name="exemple" author="totoro">
  <!-- description des connaissances -->
  <knowledge>
    <schemas> <!-- description des schémas -->
      <schema id="1" label="Etres Vivants">
        <representationDescriptionSchema id="1" x="0" y="0"
backgroundColor="00FF00"/>
      </schema>
      <schema id="2" label="Vertebres">
        <representationDescriptionSchema id="1"
backgroundColor="FF952B"/>
      </schema>
```

```

<!--voir en annexe pour la partie manquante -->

</schemas>

<relations>
  <relation name="isA" > <!-- description des relations -->
    <tuple><idSchema>1</idSchema> <idSchema>2</idSchema></tuple>
    <tuple><idSchema>3</idSchema> <idSchema>9</idSchema></tuple>
    <!--voir en annexe pour la partie manquante -->
  </relation>
</relations>
</knowledge>
<cartographies>
  <!-- description de la cartographie -->
  <cartography height="1000" width="1000">
    <representationOperations>
      <!--opération pour créer l'espace représenté -->
      <represent idSchemaRepresentationDescription="1"/>
      <representationTransformationOperations>
        <!--opération pour répartir radialement les structures visuelles -->
        <mapSchemaInStar idSchema="1" length="100"/>
        </representationTransformationOperations>
        <visualisationOperations>
          <!--opération pour visualiser l'espace représenté avec l'outils TreeBolic -->
          <visualisationInTreeBolic/>
        </visualisationOperations>
      </cartography>
    </cartographies>
  </map>

```

b. La carte

Voici le résultat obtenu avec Os Map :

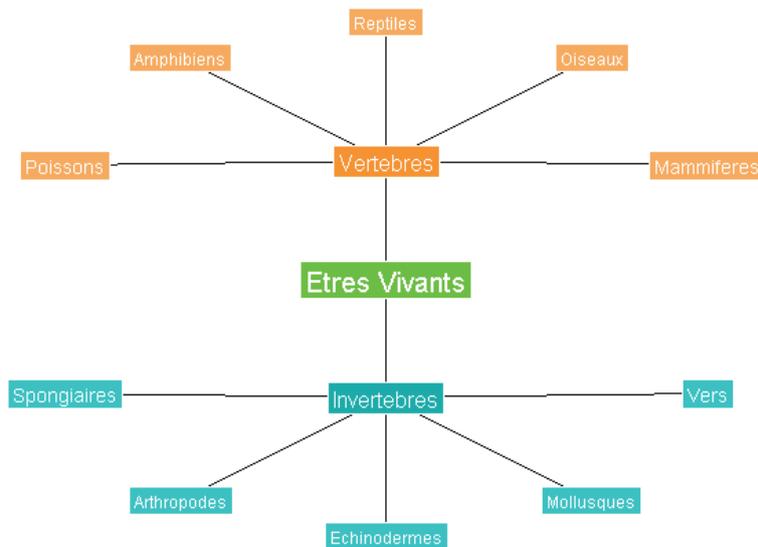


Figure 87 - Exemple de carte avec MDL.

3.5.5 Synthèse

MDL est un langage de la famille XML pour décrire une cartographie sémantique (entendue comme l'ensemble des opérations nécessaires à appliquer à un espace informationnel *structuré* pour créer des cartes). Il met directement en œuvre notre processus ainsi que notre formalisme de description SNDF. Il permet de décrire les connaissances (l'espace informationnel) à cartographier puis, la succession des opérations à leur appliquer pour aboutir aux différentes cartes.

MDL est utilisé par le système Os Map, notre service Web de cartographie sémantique, décrit ultérieurement (voir en page 163). Avec Os Map, il est possible de commander une cartographie par service web et d'obtenir la carte via Internet ; les requêtes sont alors soumises en MDL.

3.6 Notre architecture à agents et MVC distribué

Au fil des années les systèmes d'information des organisations ont fortement muté. Ils sont passés pour la plupart de systèmes monolithiques à des systèmes de plus en plus répartis. L'évolution majeure ces dernières années dans les systèmes d'information est l'arrivée des services web qui permettent aux entreprises d'offrir leurs services comme des briques logicielles interconnectables. Chaque brique logicielle est assimilable à un agent logiciel évoluant sur Internet. Cette évolution permet avant tout de répondre à des besoins internes, mais elle permet aussi aux entreprises de sous-traiter certains de leurs traitements.

La cartographie sémantique (entendue comme processus de gestion des connaissances) doit aussi être proposée comme un service pour intégrer les nouvelles architectures des systèmes d'information.

Face à cette problématique, notre objectif est de définir une nouvelle architecture dont le principe fondamental est la décomposition de la cartographie en différents services fournis par des agents logiciels. Notre architecture de cartographie sémantique est alors un ensemble d'agents qui interagissent pour produire des cartes.

3.6.1 Cartographie sémantique orientée service

Dans les organisations, la plupart des systèmes d'information sont distribués : des serveurs d'applications fournissent un ou plusieurs services et les utilisateurs utilisent un poste client pour accéder à ces services. Pour répondre aux besoins de distribution des traitements, la cartographie sémantique doit évoluer vers une architecture orientée service.

L'approche courante consiste à appréhender la cartographie de données abstraites comme un processus séquentiel monolithique assimilable à une boîte noire permettant de passer d'un ensemble de données à une carte. Cette vision contraint à exécuter le processus complet de cartographie par une seule application exécutée sur le poste client. Une telle architecture a de nombreux inconvénients. Tout d'abord, il est nécessaire de transférer l'ensemble des données à cartographier sur le poste client ; la conséquence directe est une consommation importante de bande passante entre le serveur et le poste client. Ensuite, cette architecture impose aux applications clientes de réaliser l'ensemble des opérations liées à la cartographie et elles sont généralement très gourmandes en ressources. Le poste client doit alors être performant. De plus, toute modification dans le processus de cartographie implique que tous les postes clients doivent être mis à jour.

Finalement, avec cette approche tous les postes clients doivent être performants pour effectuer les mêmes tâches (la cartographie), le serveur de services joue le rôle de serveurs de données et enfin, le réseau doit être rapide pour transférer de grandes quantités de données.

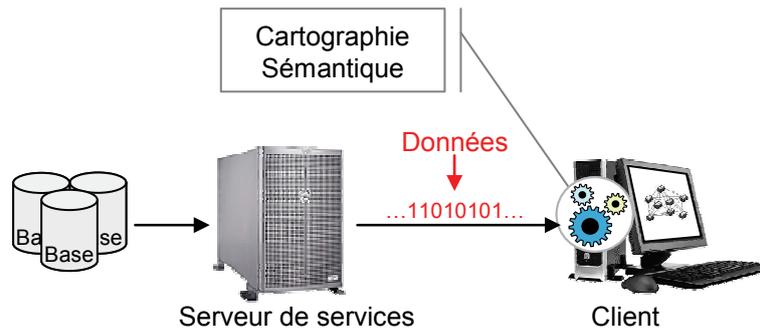


Figure 88 - Architecture courante pour la cartographie sémantique.

L'objectif de notre proposition d'architecture est de fournir un service de haut niveau pour rendre la cartographie sémantique transparente pour les applications qui souhaite proposer des cartes sémantiques.

Dans ce dessein, nous proposons de déplacer la cartographie sémantique au même niveau que les autres services web de gestion des connaissances. En d'autres termes, nous proposons une architecture à agents pour offrir des cartes sémantiques interactives directement aux applications clientes.

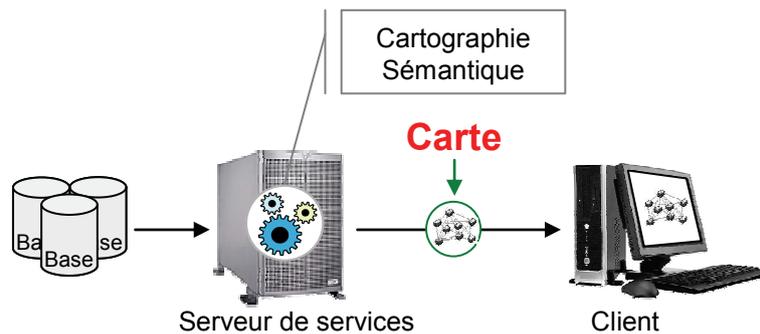


Figure 89 - Principe de notre proposition d'architecture de cartographie sémantique.

Notre proposition de processus permet aisément d'avoir une approche orientée service du processus de cartographie sémantique. L'objectif final de notre proposition d'architecture est double. D'une part, elle doit offrir un service de cartographie pour produire des cartes et d'autre part, elle doit fournir aux utilisateurs et à leurs applications les moyens pour y accéder et interagir avec elles.

3.6.2 Distribution du processus de cartographie

La carte constitue l'interface entre un espace informationnel et l'utilisateur. Elle met en œuvre des mécanismes d'interface homme-machine. Pour atteindre notre objectif, nous avons choisi le paradigme MVC qui permet de gérer les interactions entre un modèle et ses vues.

a. Model View Controller

MVC (pour « Model View Controller ») est un paradigme de spécification et de réalisation d'interfaces utilisateur [Krasner & Pope, 1988]. Sa première mise en œuvre a été effectuée

dans l'environnement de programmation Smalltalk-80. MVC est une modélisation d'environnements orientés objets permettant aux objets d'avoir plusieurs vues. L'intérêt de ce modèle est qu'il permet de maintenir la cohérence entre les différentes vues d'un ensemble d'objets.

Le principe fondamental de MVC est de dissocier le modèle, contenant les informations, de ses vues. La césure entre le monde des objets et le monde de leurs visualisations nécessite un lien entre l'objet et l'ensemble de ses vues. Dans MVC, ce lien n'est pas direct, mais il existe au travers d'une entité intermédiaire nommée « contrôleur ». Chaque vue possède un contrôleur et toutes les interactions sur la vue passent par ce contrôleur. Ce dernier est alors en charge de répercuter (ou non) les modifications de la vue jusqu'au modèle. Le modèle peut ensuite envoyer un message de mise à jour à l'ensemble de ses objets (mécanisme des « objets dépendants » de Smalltalk-80). Chaque modèle, possède un ensemble de couples constitués d'une vue et d'un contrôleur.

Pour illustrer ces interactions, prenons l'exemple d'un champ de saisie d'un entier dans une interface d'une application quelconque. Le déplacement du curseur dans ce champ de saisie n'a (par défaut) aucune répercussion sur le modèle. Dans ce cas, le contrôleur se contente de déplacer le curseur dans le champ (la vue). En revanche, si l'utilisateur saisit une valeur dans le champ (une vue), le contrôleur envoie un message au modèle. Celui-ci modifie sa valeur puis envoie au besoin un message à l'ensemble de ses vues (via le contrôleur) pour leur indiquer de se redessiner.

Le schéma suivant représente les différents liens entre le modèle et ses vues. Les interactions sont des envois de messages.

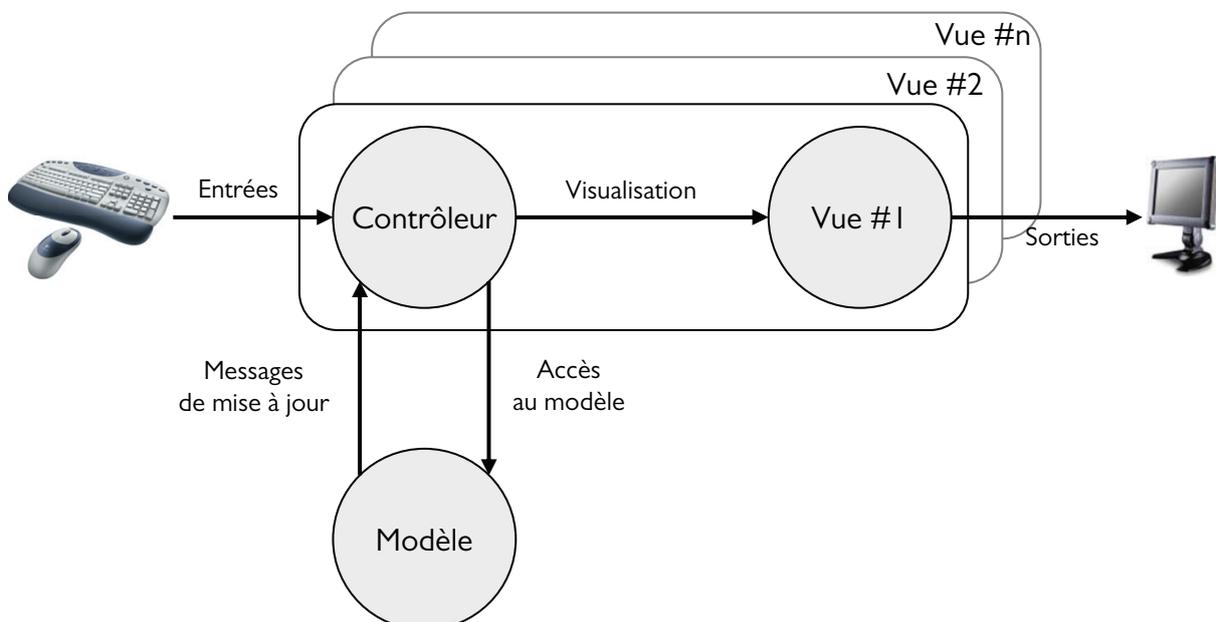


Figure 90 - Principe de MVC.

Ce mécanisme d'interactions entre le modèle et ses vues répond à nos besoins pour créer un service web de cartographie sémantique : les espaces informationnels *structurés* correspondent aux modèles et les cartes constituent les vues de ces modèles.

Pour illustrer l'utilisation de MVC pour la cartographie sémantique, prenons l'exemple de l'organigramme. Imaginons que la carte serve d'interface à l'utilisateur pour effectuer un

ensemble d'opérations sur chaque collaborateur. Pour notre exemple imaginons deux opérations : l'affichage des informations détaillées de l'utilisateur sélectionné dans un panneau latéral et la suppression d'un collaborateur. Ces deux opérations sont accessibles à l'aide d'un menu contextuel. Ainsi, lorsque l'utilisateur sélectionne un collaborateur (via un contrôleur), un menu contextuel s'affiche et permet de déclencher l'une des opérations. Si l'opération d'accès aux informations détaillées est sélectionnée, le contrôleur de la carte envoie un message à l'application qui met à jour le panneau latéral (correspondant à une carte des informations détaillées). En revanche, si l'utilisateur choisit de supprimer un collaborateur, le contrôleur envoie un message de modification à l'espace informationnel et envoie un message de mise à jour de toutes les cartes qui à leur tour déroulent à nouveau leur processus de cartographie.

Le mécanisme des envois de messages est un moyen simple et robuste pour maintenir la cohérence entre les vues et leur modèle. Cependant il est limité lorsque l'on essaye d'appliquer MVC à une architecture distribuée.

b. Distribution de MVC et système à agents

Pour mettre en œuvre une architecture orientée service de cartographie sémantique, nous avons besoin de distribuer MVC avec les cartes sur les postes clients (la vue et son contrôleur) et leur modèle sur le serveur de services (l'espace informationnel).

Ces travaux ont été réalisés dans un contexte industriel particulier (voir à la page 159) ce qui nous contraint d'utiliser des standards pour concevoir notre architecture. La plus forte était d'utiliser les services web via une architecture de type « client/serveur ».

Dans notre cas, on peut schématiser cette architecture par la figure suivante :

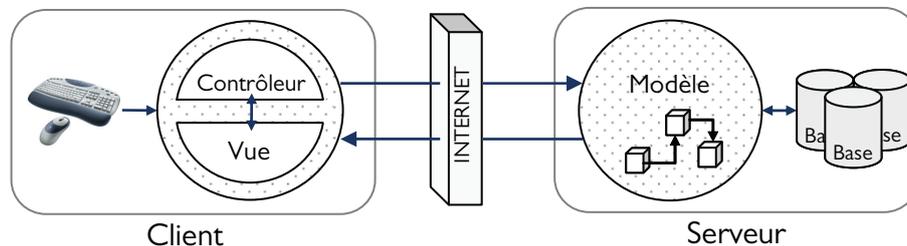


Figure 91 - MVC : distribution du modèle et de ses vues.

MVC implique des communications allant des contrôleurs vers leur modèle (communication de type « modification ») et dans le sens inverse, des modèles vers leurs vues (communication de type « mise à jour »). La principale difficulté réside dans le fait que MVC n'a pas été conçu pour supporter la distribution des modèles et de leurs vues sur des machines physiques différentes avec une architecture client/serveur.

Dans une architecture Internet, les communications dans le sens du client vers le serveur ne posent aucune difficulté car ce sont des requêtes « classiques » : le client demande une information au serveur. Dans notre cas, chaque couple (vue, contrôleur) peut, sur la base de protocoles Internet, envoyer des messages aux modèles. En revanche pour des raisons de sécurité, les communications dans le sens inverse (du serveur vers le client) ne sont pas autorisées : le serveur ne peut pas envoyer d'informations aux clients sans sollicitation préalable de leur part.

Dans notre contexte de cartes distribuées par rapport à leur modèle (l'espace informationnel), les messages spontanés du modèle vers les vues sont impossibles. Il ne peut donc pas y avoir de demandes de mise à jour de la part d'un modèle pour l'ensemble de ses vues. Nous sommes donc contraint de limiter la capacité des systèmes à agents logiciels pour s'adapter à une architecture de type client/serveur. Plus particulièrement, il est nécessaire de trouver une solution pour informer les vues des différents changements du modèle et il reviendra alors aux vues de se mettre à jour. Cette solution revient à adapter le mécanisme des « objets dépendants » de Smalltalk-80 garantissant ainsi la cohérence globale.

Notre proposition d'architecture définit une extension du modèle MVC pour obtenir un modèle MVC distribué. De plus, cette architecture permet (à l'instar du modèle original) de travailler simultanément sur plusieurs cartographies d'un même espace informationnel constituant autant de modèles cartographiés.

Pour parvenir à notre objectif, nous avons défini une nouvelle architecture dont le principe fondamental est la décomposition de la cartographie en différents services fournis par des agents logiciels [Woolridge, 2001]. Notre service web de cartographie sémantique est alors assimilable à un ensemble d'agents qui interagissent pour produire une carte.

Dans le cas de cette architecture, l'espace informationnel *brut* correspond aux différentes bases possibles des organisations (comme des bases de données ou des bases de connaissances). Pour simplifier la lecture nous désignerons l'espace informationnel *brut* par le terme « base ».

Nous avons décomposé notre architecture selon les principes de notre extension de MVC ce qui nous a permis de définir cinq types d'agents logiciels. Ces types d'agents sont présentés dans le tableau suivant :

Type	Rôle
 <p data-bbox="252 1451 432 1563">Service de Cartographie Sémantique</p>	<p data-bbox="531 1189 1414 1294">Le « Service de Cartographie Sémantique » est chargé de recevoir les requêtes de cartographie des clients. Il est hébergé par le serveur de services.</p> <p data-bbox="531 1317 1414 1386">Pour chaque requête de cartographie, il crée un agent Cartographe en charge de gérer toute la cartographie.</p> <p data-bbox="531 1408 1414 1514">Il est aussi en charge de router les messages entre les gestionnaires de cartes (côté client) et les différents cartographes (côté serveur).</p> <p data-bbox="531 1536 1414 1606">Si un client n'exploite plus ses cartes, il est en charge de supprimer l'agent Cartographe associé.</p>

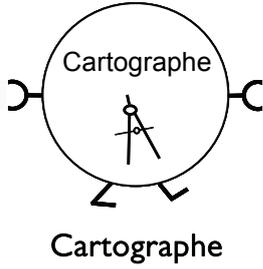
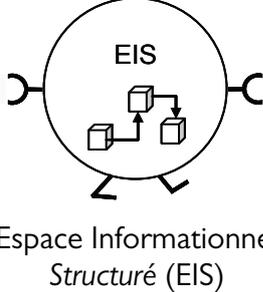
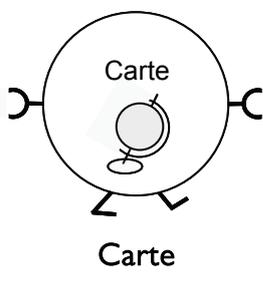
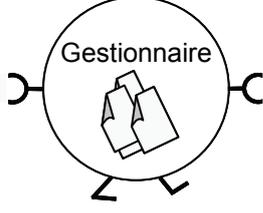
	<p>L'agent « Cartographe » est créé suite à une demande de cartographie au service de cartographie sémantique. Il est hébergé par le serveur et il est associé à un seul client.</p> <p>Il est chargé de gérer le processus complet de cartographie en créant les espaces informationnels <i>structurés</i> (le modèle en MVC) associés aux cartes puis, en déroulant toutes les opérations de cartographie sur le modèle.</p> <p>Une cartographie peut contenir plusieurs cartographies ce qui implique que l'agent Cartographe crée autant d'agents d'Espace Informationnel <i>Structuré</i> (EIS) que de cartographies.</p> <p>L'agent Cartographe est aussi responsable de toutes les interactions entre l'utilisateur et les EIS ainsi que de maintenir la cohérence entre les différentes cartes.</p>
	<p>L'agent « Espace Informationnel <i>Structuré</i> » (EIS) est créé par un agent Cartographe à raison d'un agent EIS par demande de cartographie. Il offre une structuration de l'espace informationnel et il est hébergé par le serveur.</p> <p>Il a pour fonction de gérer l'espace informationnel <i>structuré</i> associé à une ou plusieurs cartes. Pour construire cet espace, il puise dans la base en fonction de la requête de l'utilisateur. Par exemple, il récupère toutes les entités de la base liées par la relation de généralisation/spécialisation.</p>
	<p>L'agent « Carte » est l'agent en charge du couple vue et contrôleur. Il supporte les interactions entre l'utilisateur et la carte.</p> <p>Il gère les interactions qui portent sur la représentation (zoom, déplacement, sélections, etc.). Pour les interactions qui portent sur les connaissances cartographiées (modification, suppression, etc.) il prévient l'agent EIS de la carte via l'agent gestionnaire de cartes.</p> <p>Une carte est associée à un seul EIS, mais deux cartes peuvent avoir le même EIS. Cet agent offre donc un point de vue sur l'espace informationnel et il est hébergé par le client.</p>
	<p>L'agent « Gestionnaire de Cartes » est en charge de la coordination entre les différentes cartes et leur agent EIS.</p> <p>Tout comme le service de cartographie, il est en charge de router les messages entre le cartographe (côté serveur) et ses cartes (côté client).</p> <p>Il est hébergé sur le client.</p>

Tableau 7 - Agents logiciels de notre architecture.

Le schéma ci-dessous représente un cas de figure possible de notre architecture : deux espaces informationnels *structurés* sont cartographiés en trois cartes. Ces deux modèles constituent deux modélisations possibles de l'espace informationnel *brut* (les bases) et les cartes en constituent trois points de vue différents. Le schéma représente aussi les différentes interactions entre les agents. Chaque agent Carte à la même couleur que l'agent EIS associé.

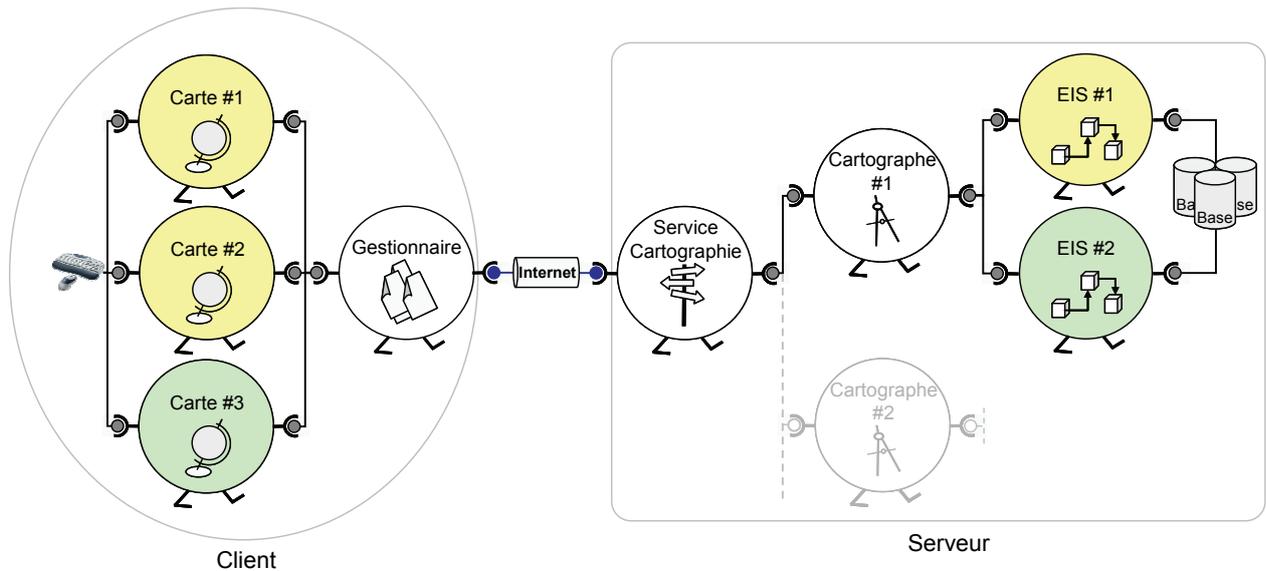


Figure 92 - Configuration possible de notre architecture de cartographie sémantique distribuée.

c. Mécanisme des envois de messages

Il existe deux modes de communications entre les agents selon qu'ils évoluent dans le même système ou non (poste client ou service de cartographie).

Le premier cas, qui est le plus simple, est celui de deux agents appartenant au même système. Les communications sont alors dites « classiques » (assimilables à des appels de méthode en programmation objet). Par exemple, un agent Carte demande à son agent Gestionnaire de mettre à jour l'EIS par un appel à une méthode particulière en passant l'élément à modifier ainsi que sa nouvelle valeur en paramètres.

Le deuxième cas, qui est quant à lui beaucoup plus complexe, est celui d'une communication entre des agents distribués sur des systèmes différents. Les communications sont alors véhiculées sur un médium comme Internet. C'est dans ce cas qu'interviennent les difficultés liées à l'architecture client/serveur. Nous avons vu que les communications dans le sens du client vers le serveur sont simples. En revanche, celles qui s'effectuent dans le sens inverse sont impossibles spontanément mais, elles sont nécessaires pour garantir la cohérence entre chaque modèle et ses vues.

Nous avons pu trouver une solution à ce problème suite à une observation importante sur MVC appliqué à la cartographie sémantique : un modèle ne demande pas *spontanément* de mettre à jour l'ensemble de ses vues. Il le fait uniquement après une modification du modèle par une des vues. Nous avons donc exploité cette particularité avec l'agent de gestion de cartes et le cartographe. En effet, lorsque l'utilisateur modifie un EIS à l'aide d'une de ses cartes, une requête de mise à jour est envoyée à l'agent cartographe. Dans notre architecture, ce type de requête est constitué de deux éléments : l'EIS concerné (lié à la vue) et la mise à jour nécessaire (par exemple l'entité à supprimer).

Lorsque l'agent Cartographe reçoit un message de modification, il le redirige vers l'agent EIS concerné. Ce dernier modifie son espace informationnel et si besoin modifie la base. Si la base a été modifiée, l'agent EIS avertit le Cartographe qu'il faut mettre à jour tous les EIS et à son tour, il prévient les agents EIS de se mettre à jour. Dans tous les cas, l'agent EIS avertit en retour l'agent Cartographe des vues qui doivent être mises à jour. À son tour, l'agent

Cartographe retourne à l'agent Gestionnaire la liste des vues à mettre à jour. Ce dernier étant sur le même système que les cartes, il peut prévenir les agents de Cartes concernés. Cette liste des cartes à mettre à jour n'est alors que le retour du message envoyé par le client. Il n'y a alors eu qu'une communication entre les deux systèmes (client et serveur) : la requête et son résultat.

Le schéma suivant est un exemple de diagramme de séquences simplifié pour illustrer les communications entre les différents agents. Cet exemple est basé sur deux cartes issues de deux espaces informationnels *structurés* différents provenant d'une même base. L'utilisateur supprime une entité « e1 » de la base via l'agent Carte #1 :

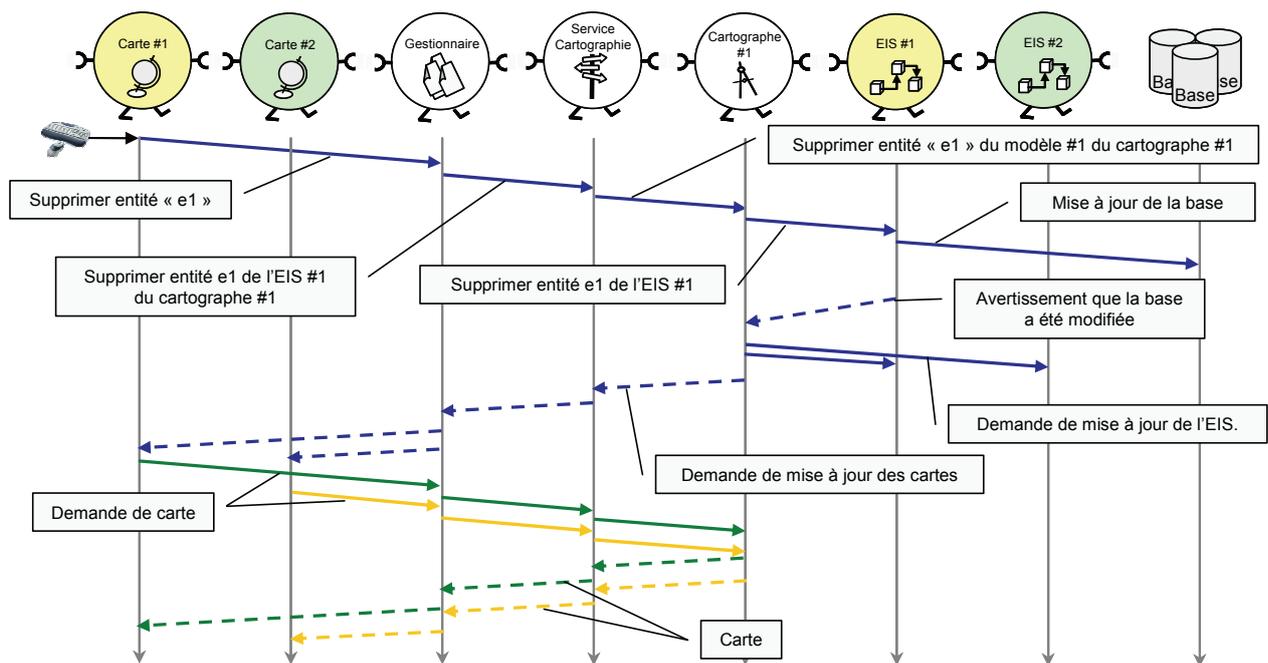


Figure 93 - Mécanisme des communications : illustration avec un diagramme de séquences.

Ce mécanisme de communication entre les agents permet de garantir la cohérence entre les cartes et leur espace informationnel. Néanmoins, il reste une difficulté pour garantir la cohérence en plusieurs espaces informationnels (agents EIS) qui partagent simultanément une même base. En effet si un utilisateur modifie la base via son agent Cartographe, les autres agents Cartographes (correspondant aux autres utilisateurs) seront avertis de cette modification. Le problème survient s'il leur est aussi nécessaire de modifier leurs cartes car ils ne peuvent pas avertir spontanément leurs agents gestionnaires. Pour contourner ce problème, nous avons spécifié aux agents Gestionnaire de demander régulièrement (de l'ordre de la minute) à leur agent Cartographe si des mises à jour sont nécessaires.

3.6.3 Synthèse

Nous avons défini une nouvelle architecture dont le principe fondamental est la décomposition de la cartographie en différents services fournis par des agents logiciels. Notre architecture de cartographie sémantique est alors un système d'agents logiciels qui interagissent pour produire des cartes.

L'objectif final de notre proposition d'architecture est de construire une carte et de fournir à l'utilisateur les moyens pour y accéder et interagir avec celle-ci. La carte constitue l'interface entre l'espace informationnel et l'utilisateur. Pour permettre ces interactions, il faut mettre en œuvre des mécanismes d'interface homme-machine. Pour notre proposition, nous avons choisi le paradigme MVC qui permet de gérer des interactions entre un espace informationnel et ses vues. De plus, notre proposition d'architecture définit une extension du modèle MVC pour obtenir un modèle MVC distribué. Cette évolution permet (à l'instar de l'original) de travailler simultanément sur plusieurs cartographies de sous-ensembles d'une même base constituant autant d'espaces informationnels cartographiés.

Finalement, nous avons décomposé notre architecture selon les principes de notre extension de MVC ce qui nous a permis de définir cinq types d'agents logiciels. Ces différents agents interagissent entre eux pour construire des cartes selon notre approche de cartographie sémantique.

Cette architecture est à la base d'Os Map, notre service web de cartographie sémantique que nous avons mis en œuvre. Il est décrit dans la partie qui présente nos réalisations (voir à la page 163).

3.7 Synthèse des propositions

Si différentes qu'elles puissent paraître, nos propositions constituent un tout cohérent, c'est-à-dire une **méthode** pour cartographier un espace informationnel selon nos besoins. Elle prend en compte la sémantique du domaine et exploite les retours d'expérience des utilisateurs.

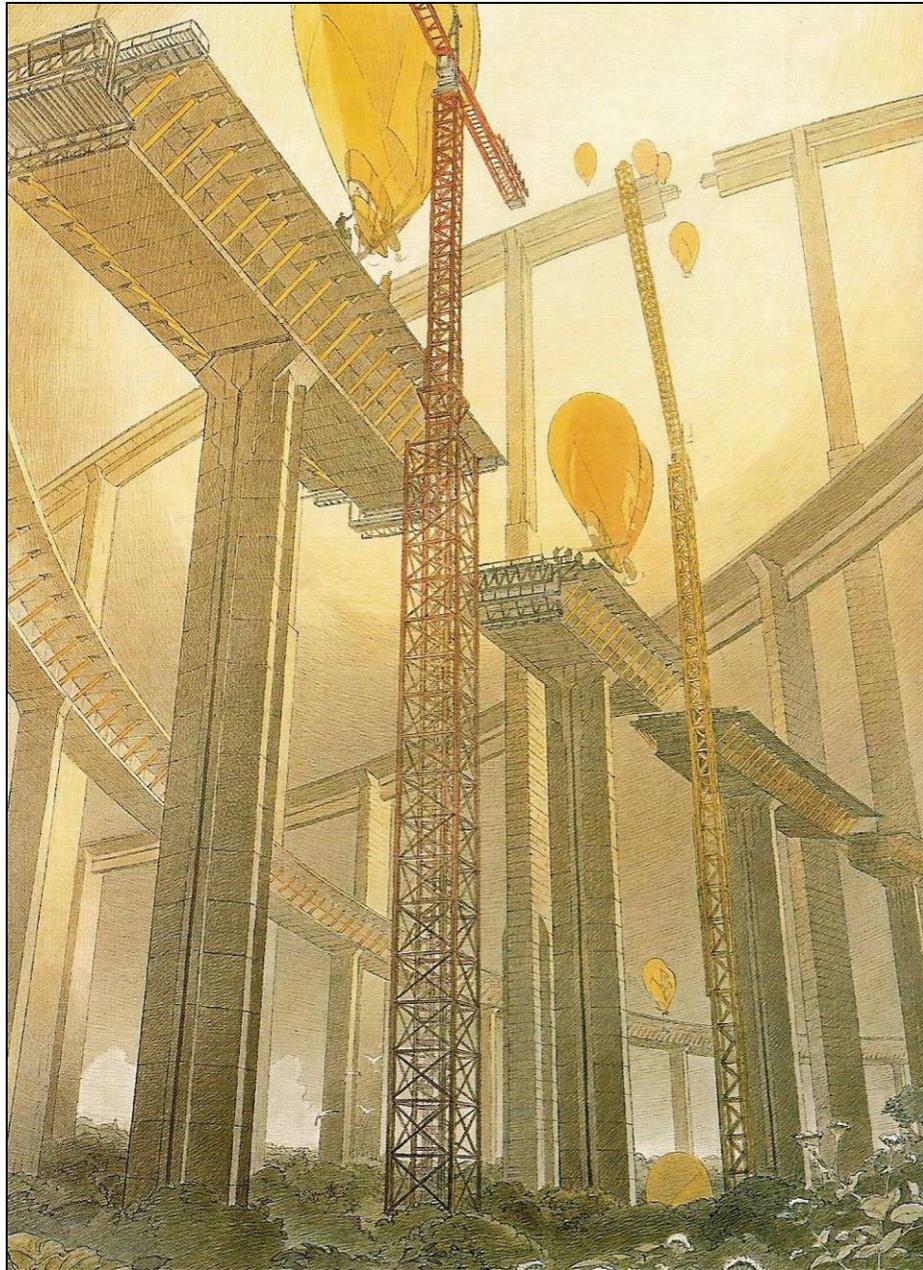
Elle repose sur un nouveau **processus** qui assimile la cartographie à un ensemble d'opérations permettant de construire une carte d'un espace informationnel *brut* en passant successivement par un espace informationnel *structuré* puis un espace informationnel *représenté* et enfin un espace informationnel *visualisé*. La mise en œuvre de ce processus a requis la création d'un **formalisme de description** dédié à la cartographie sémantique que nous avons nommé SNDF (Semantic Network Description Formalism).

Nous avons aussi proposé des **critères pour la cartographie** d'espaces informationnels *structurés* par une ontologie de domaine (comme l'utilisation de techniques « focus+context » avec une géométrie euclidienne). La prise en compte de ces critères a abouti à la spécification de deux nouveaux paradigmes pour explorer un espace informationnel *structuré* : l'EyeTree et le RadialTree.

La cartographie sémantique a aussi requis la définition d'un **langage** de haut niveau, nommé MDL (Map Description Language), pour construire une carte en décrivant sa cartographie sémantique. Enfin, nous avons proposé une **architecture** de cartographie sémantique basée sur une extension du modèle MVC (Model View Controller). Elle permet d'intégrer la cartographie sémantique dans les solutions de gestion de connaissances à l'aide d'un ensemble d'agents logiciels.

La partie suivante présente nos différentes réalisations associées à cette étude. Elles sont toutes la mise en œuvre directe de nos différentes propositions.

Partie 4 : **Réalisations des propositions**



« Il faut toujours garder à l'esprit que réaliser une carte n'est affaire que de cohérence et de respect des subtilités de la vision humaine. La cartographie, à la fois, art et science, à la jonction de la technologie, de la communication et du graphisme, déploiera alors toutes ses qualités d'aide à l'action. »

Didier Poidevin [Poidevin, 1999]

Rappel du plan

4.1 Contexte.....	159
4.1.1 Équipe de Recherche Technologique	159
4.1.2 La démarche Ousia	160
4.2 Cartographie Sémantique avec Os Map.....	163
4.2.1 Objectif et principe	163
4.2.2 Services web et architectures orientées services	164
4.2.3 L'évocation du service	166
4.2.4 La mise en œuvre.....	167
4.2.5 Exemples	169
4.2.6 Synthèse.....	171
4.3 Modélisation graphique avec SNCW	172
4.3.1 La modélisation graphique de connaissances	173
4.3.2 Besoins pour la modélisation graphique	176
4.3.3 Le principe de SNCW	178
4.3.4 L'architecture de SNCW.....	182
4.3.5 Intégration dans KCW	188
4.3.6 Synthèse.....	188
4.4 Exploration sémantique avec EyeTree et RadialTree.....	189
4.4.1 Os EyeTree	189
4.4.2 Os RadialTree	192
4.4.3 Synthèse.....	194
4.5 Synthèse des réalisations	195

Cette partie présente les associées à cette étude. Elles ont été conçues dans le contexte particulier d'une Équipe de Recherche Technologie (ERT) associée à une société de valorisation leur conférant alors un caractère très industriel.

La première réalisation présentée est Os Map, notre service web de cartographie sémantique. Ce service met directement en œuvre l'ensemble de nos propositions (méthode, processus, formalisme de description, paradigmes, langage et architecture).

SNCW (Semantic Network Craft Workbench) correspond à notre deuxième réalisation ; il permet d'exploiter notre méthode de cartographie sémantique pour construire graphiquement une modélisation pour un domaine donné. SNCW met aussi en œuvre SNDF, notre proposition de formalisme ainsi que notre processus.

Enfin, les deux réalisations présentées sont l'EyeTree et le RadialTree correspondant à l'implémentation de notre spécification de paradigmes pour l'exploration.

4.1 Contexte

Les parties qui suivent présentent les différentes réalisations associées à ces travaux. Elles ont été réalisées dans le cadre d'une équipe de recherche (Équipe Condillac) reconnue ERT (Équipe de Recherche Technologique).

4.1.1 Équipe de Recherche Technologique

« La recherche technologique a pour finalité de résoudre des problèmes posés par des entreprises industrielles ou de service ou liés à une demande sociétale. Elle se justifie par une demande du milieu socio-économique et ses thèmes sont donc inspirés par des objets industriels, produits, procédés, systèmes génériques ainsi que des services ». « Une équipe de recherche technologique (ERT) est une équipe de recherche qui, en partenariat avec des industriels, mène sur le moyen terme des recherches dans le cadre de projets visant à lever des verrous technologiques relatifs à des problèmes qui n'ont pas de solutions immédiates. Pour être reconnue, une ERT doit s'appuyer sur une recherche amont de qualité et justifier d'un engagement fort d'industriel(s) »¹. Par conséquent, une ERT est contrainte de réaliser d'une part, une recherche technologique pour lever des verrous technologiques et d'autre part, de valoriser ses travaux dans le cadre d'un partenariat industriel.

Les travaux associés à cette étude ont été réalisés au sein de l'équipe Condillac qui a été reconnue ERT par le Ministère de la Recherche. Le partenariat industriel s'effectue avec l'entreprise Ontologos Corp. qui s'engage à valoriser l'ensemble des travaux de recherche.

Les travaux dans une ERT sont « bipolaires » avec un pôle recherche (l'équipe Condillac) et un pôle industriel (l'entreprise Ontologos Corp.). Ils sont en prise directe avec le monde des entreprises. Ce contexte d'ERT apporte des contraintes liées au monde industriel. Il en résulte des réalisations qui sont d'une part, le fruit du travail de recherche (pour lever des verrous technologiques) et d'autre part la réponse directe à de réels besoins (exprimés par des organisations par l'intermédiaire de l'entreprise en charge de la valorisation).

¹ <http://www.recherche.gouv.fr/technologie/rechped/default.htm>

La cartographie sémantique fait partie des verrous technologiques que l'équipe Condillac doit lever. Les réalisations présentées ici ont été guidées par la volonté de lever ce verrou, mais aussi par la volonté de répondre aux besoins du monde de l'entreprise.

Les retours d'expérience des différentes applications constituent alors la plus efficace des validations.

4.1.2 La démarche Ousia

Les différentes réalisations qui accompagnent ces travaux s'intègrent dans la suite de solutions proposées par l'équipe Condillac et la société Ontologos Corp. : « la démarche Ousia ».

La démarche Ousia correspond à la gestion de l'information par le sens métier. Elle met en œuvre une méthode et un ensemble d'outils pour construire et exploiter les connaissances des organisations. Le principe fondamental de cette démarche est la gestion des informations selon leur sens métier. Cette méthode repose sur la construction d'un référentiel métier propre à l'organisation. Ce référentiel est la pierre angulaire de toute la démarche.

a. La modélisation ontologique comme référentiel métier

Le référentiel métier est porteur du sens du domaine et il repose sur l'ontologie des concepts (les idées) et sur la langue (les mots pour les dire) du domaine.

Bien que les ontologies aient une visée normative, il existe de nombreuses définitions et approches différentes, parfois contradictoires. La définition que nous utilisons synthétise la plupart des définitions existantes :

« Définie pour un objectif donné et un domaine particulier, une ontologie est pour l'ingénierie des connaissances une représentation d'une modélisation d'un domaine partagée par une communauté d'acteurs. Objet informatique défini à l'aide d'un formalisme de représentation, elle se compose principalement d'un ensemble de concepts définis en compréhension, de relations et de propriétés logiques »
[Roche, 2005].

La démarche Ousia permet de modéliser ontologiquement les connaissances des organisations en construisant leurs référentiels métiers. Elle repose sur des principes pluridisciplinaires (épistémologie, linguistique, intelligence artificielle et logique), sur le type d'applications visées et sur les objectifs à atteindre (consensus, cohérence, partage et réutilisabilité) [Roche & Million-Rousseau, 2003].

À l'aide de ces référentiels, il devient possible de traiter l'ensemble des informations du domaine. Prenons l'exemple de la gestion documentaire : le référentiel métier permet tout d'abord, d'indexer l'ensemble des documents de l'organisation sur l'ontologie du domaine (associer les documents aux concepts traités) et par la suite, de fournir aux collaborateurs l'information nécessaire à leur activité.

La construction de référentiels est un travail long et difficile souvent identifié comme étant le frein principal à la mise en œuvre d'une solution efficace de gestion des connaissances. C'est pourquoi la méthode Ousia propose une méthode outillée pour la construction et la maintenance de ces référentiels qui repose sur le traitement linguistique et ontologique des informations des organisations. L'analyse linguistique des documents existants permet d'extraire des lexiques de candidats termes soumis à la validation des experts. L'étude de la

structure des lexiques permet d'identifier des catégories de concepts ainsi que certains caractères définitionnels et différenciateurs qui participent à la définition des concepts de l'ontologie.

b. L'exploitation des référentiels métiers

La méthode Ousia se déroule alors en deux étapes : la modélisation puis l'exploitation.

La première étape est la modélisation des connaissances du domaine pour produire un référentiel métier. Pour identifier et décrire les concepts manipulés par le domaine, la méthode traite l'ensemble des informations véhiculées par l'entreprise, généralement des documents scientifiques et techniques. Cette étape est supportée par des ateliers logiciels (les « Craft Workbench »). Ils sont utilisés par des ingénieurs cognitivistes (« ontologue » et « terminologue ») aidés des experts du domaine.

La seconde étape est l'exploitation des référentiels construits pour traiter l'ensemble des informations de l'organisation. Cette étape est supportée par un ensemble d'outils (leur nom est préfixé par « Os »). Ces derniers sont utilisés par les clients.

c. KCW

La démarche Ousia est supportée par un ensemble d'outils qui définissent ce que nous nommons « KCW » pour Knowledge Craft Workbench : notre support (méthode et outils) de gestion des connaissances.

Le schéma ci-dessous représente les principaux éléments logiciels de KCW :

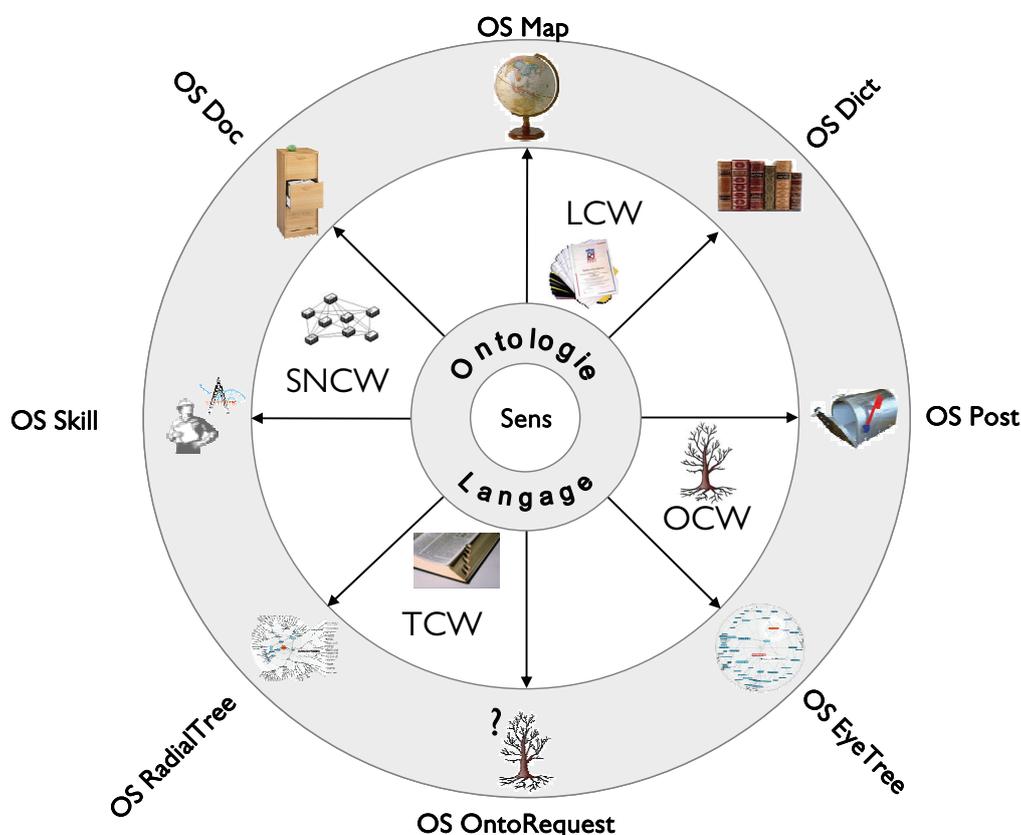


Figure 94 - KCW et la démarche Ousia.

Dans cette figure, les éléments de KCW sont représentés sur des cercles concentriques avec :

- > Au centre, le référentiel. Il est supporté par l'ontologie et le langage et il permet d'accéder au sens métier.
- > Sur le disque suivant, les ateliers de construction du référentiel :
 - **OCW** : **O**ntology **C**raft **W**orkbench est un atelier de construction de la partie ontologique du référentiel.
 - **LCW** : **L**inguistic **C**raft **W**orkbench est un atelier de traitement des ressources linguistique de l'organisation.
 - **TCW** : **T**erminology **C**raft **W**orkbench est un atelier de construction de la partie terminologique du référentiel.
 - **SNCW** : **S**emantic **N**etwork **C**raft **W**orkbench est un atelier de construction de réseaux sémantiques. Cet outil est une des réalisations associées à cette étude.
- > Sur le dernier disque, on retrouve les outils permettant d'exploiter le référentiel à des fins de gestion de connaissances :
 - **Os Map** : service web de cartographie sémantique. Il s'agit de la première réalisation qui accompagne cette étude et elle est décrite dans cette partie de ce document.
 - **Os EyeTree**, **Os RadialTree** et **Os OntoRequest** : ce sont des paradigmes de cartographie de bases de connaissances structurées ontologiquement.
 - **Os Doc** et **Os Dict** : applications de gestion de la documentation des organisations (GED).
 - **Os Skill** : application Internet/intranet de gestion des compétences.
 - **Os Post** : application de routage de mails.

Ce qui est important de retenir de cette courte présentation est qu'il existe deux grandes familles d'éléments logiciels : les ateliers pour la construction et les outils pour l'exploitation.

Les réalisations associées à cette étude s'intègrent donc pleinement dans KCW et correspondent aux besoins de la démarche Ousia.

La présentation des réalisations débute par Os Map, un service web de cartographie sémantique mettant directement en œuvre l'ensemble de nos propositions (méthode, processus, modélisation, paradigmes, langage et architecture).

SNCW correspond à notre deuxième réalisation. Il permet d'exploiter notre méthode de cartographie sémantique pour construire une modélisation d'un domaine donné. SNCW met aussi en œuvre notre proposition de modélisations ainsi que notre processus.

Enfin, les deux réalisations présentées sont l'EyeTree et le RadialTree correspondant à l'implémentation de notre spécification de paradigmes pour l'exploration.

4.2 Cartographie Sémantique avec Os Map

Au sein de l'équipe Condillac et dans le cadre de nos travaux de valorisation associés à l'entreprise Ontologos Corp., nous avons mis en œuvre notre approche orientée service pour fournir un service web de cartographie sémantique nommé « Os Map ».

Os Map est notre réalisation d'un service web de cartographie sémantique mettant directement en œuvre l'ensemble de nos propositions. Tout d'abord, il exploite notre méthode et notre processus de cartographie sémantique. Ensuite, il met en œuvre SNDF, notre formalisme de description de cartographie ainsi que notre architecture à agents logiciels. Enfin, les cartes qu'il permet de construire sont décrites avec MDL, notre langage de cartographie et elles sont visualisables avec nos deux paradigmes d'exploration, l'EyeTree et le RadialTree.

4.2.1 Objectif et principe

Os Map est notre service web de cartographie sémantique. Il a été conçu pour être intégré à toute application qui souhaite cartographier sémantiquement son espace informationnel. Le cas le plus courant est celui d'une application web (service tiers) qui souhaite intégrer des cartes dans ses différentes pages.

Concrètement, Os Map permet de cartographier toutes requêtes de cartographies soumises en MDL et de construire la carte associée. Les cartes produites sont hébergées par le système et elles sont accessibles comme ressources Internet (identifiée par un URI). Dans sa version actuelle, les cartes offertes sont disponibles sous plusieurs formats (SVG, X3D, Applet Java, images, etc.) en combinant différents paradigmes de cartographie (géométrie non euclidienne, 3D, etc.).

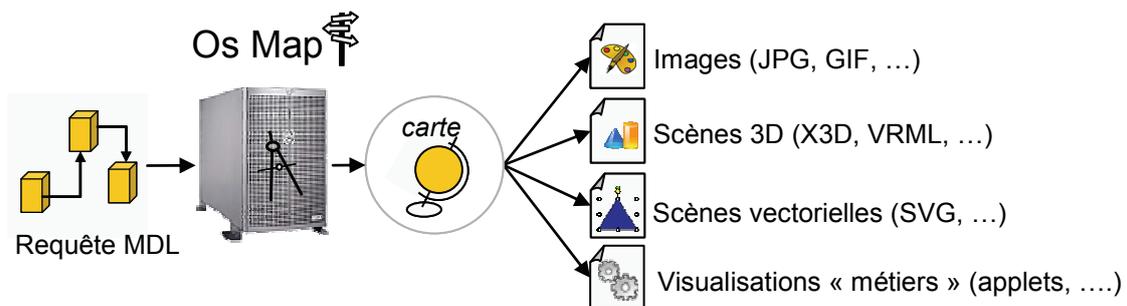


Figure 95 - Os Map construit des cartes selon différents formats.

Le rôle d'Os Map est double. Premièrement, Os Map joue le **rôle de cartographe** car il doit construire les cartes associées aux requêtes qui lui parviennent. Deuxièmement, Os Map joue le **rôle d'atlas** en hébergeant les cartes construites pour les rendre accessibles via Internet. Son rôle d'atlas est très important car il permet tout d'abord de télécharger l'application tierce. De plus, ce rôle permet d'éviter de transférer inutilement la carte entre les différents services sachant qu'elle peut être de taille relativement importante et que seul le client final a besoin de l'afficher. Et enfin, l'atlas permet de stocker les cartes pour une

consultation ultérieure par les clients sans nécessiter de reconstruire la carte à chaque utilisation.

Pour comprendre cette dualité de rôle, prenons l'exemple d'un organigramme. Imaginons une entreprise qui possède un portail Internet (l'application tierce) et qu'elle souhaite y afficher l'organigramme de ses collaborateurs. Les informations visualisées dans cet organigramme sont fonctions des droits de l'utilisateur qui la demande et du point de vue demandé. Chaque carte est alors unique et elle est cartographiée dynamiquement à chaque requête.

Lorsqu'un utilisateur demande une carte (comme l'organigramme) via une application tierce (comme le portail), ce dernier soumet une requête MDL à Os Map. Cette requête contient l'espace informationnel *structuré* à cartographier (comme les collaborateurs et leurs relations) ainsi que la liste des opérations à leur appliquer (représentation, répartition, etc.).

Pour chaque requête de cartographie, Os Map construit la carte associée et la stocke dans son atlas. Chaque carte possède un URI¹ et correspond à une page Internet contenant toutes les sous-cartographies. Cet URI est retourné comme résultat de cartographie à l'application tierce. Le client (via un navigateur Internet par exemple) n'a alors plus qu'à afficher la page associée à l'URI (ce qui a pour conséquence de récupérer la carte dans l'Atlas).

Voici un schéma illustrant le principe d'Os Map :

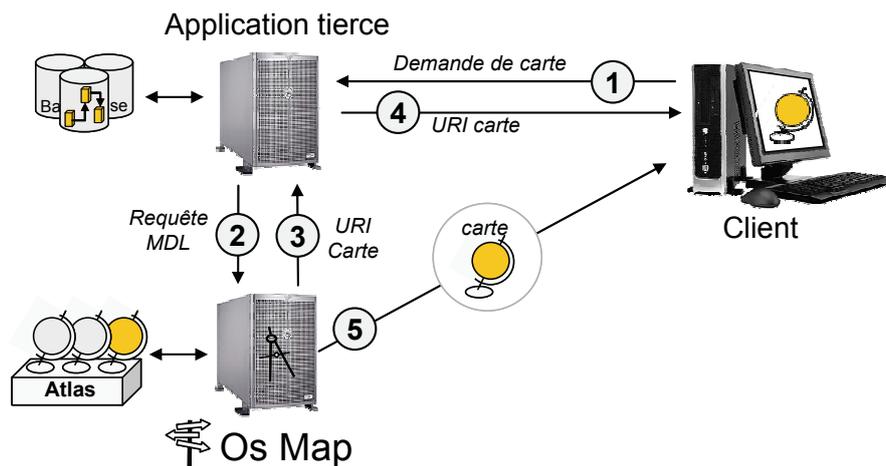


Figure 96 - Usage d'Os Map.

Notre service de cartographie sémantique est un service web. La section suivante rappelle brièvement le principe des services web.

4.2.2 Services web et architectures orientées services

Pour faire face aux décentralisations et aux besoins de sous-traitance, les organisations ont ouvert leurs réseaux internes vers Internet. De nouveaux besoins ont entraîné la création de nouvelles solutions et plus particulièrement de nouvelles technologies. Malheureusement aucune d'elles n'a fait autorité et l'unanimité auprès des entreprises.

¹ « Un URI, de l'anglais Uniform Resource Identifier, soit littéralement identifiant uniforme de ressource, est une courte chaîne de caractères identifiant une ressource physique ou abstraite, et dont la syntaxe respecte une norme d'Internet mise en place pour le World Wide Web. » <http://fr.wikipedia.org/wiki/URI>

Il y a quelques années (environ 6 ans), les services web sont apparus. D'après le W3C¹ (World Wide Web Consortium) – organisme en charge de standardiser les protocoles Internet – les services Web sont définis comme suit :

- > Un service web est un système logiciel identifié par un URI, dont les interfaces publiques et les liens sont définis et décrits en XML.
- > Sa définition peut être découverte par d'autres systèmes logiciels.
- > Ces derniers peuvent interagir avec le service web d'une manière prescrite par sa définition, en utilisant des messages XML portés par les protocoles Internet.

Depuis leur création, les services Web sont perçus comme la solution tellement attendue aussi bien par le monde des entreprises que par celui de la recherche. Les services web sont le fruit d'une longue évolution. Ils prennent leurs fondements dans des domaines de recherche beaucoup plus anciens comme l'intelligence artificielle et le génie logiciel.

Les services web ne sont que la facette technique d'une évolution plus générale nommée « SOA » ou « Architectures Orientées Services ». Les SOA ont pour objectif de rendre les systèmes d'information flexibles et agiles pour répondre aux besoins de tous les acteurs du monde économique. Concrètement, la méthode associée aux SOA permet de décloisonner les applications pour offrir de nouveaux services et une plus grande réactivité des systèmes d'information. Les SOA ne sont pas attachées à une technologie particulière, mais les services web en constituent à l'heure actuelle la base technique. Ils permettent de réaliser des SOA en distribuant sur Internet ou Intranet les applications existantes des systèmes d'information.

La notion fondamentale de cette évolution est la notion de « service » : chaque traitement est appréhendé et construit comme un service offert par un agent logiciel hébergé par un serveur. Le système d'informations devient alors un ensemble d'agents communicants. Avec les SOA, les systèmes d'information reposent sur une architecture de services (des applications) pour créer de nouvelles applications composites (un nouveau service).

Lorsqu'une application souhaite utiliser un service Web, elle lui envoie une « invocation » et ce dernier lui renvoie une « réponse ». Les services reliés peuvent être internes à l'entreprise ou proposés par des tiers par sous-traitance.

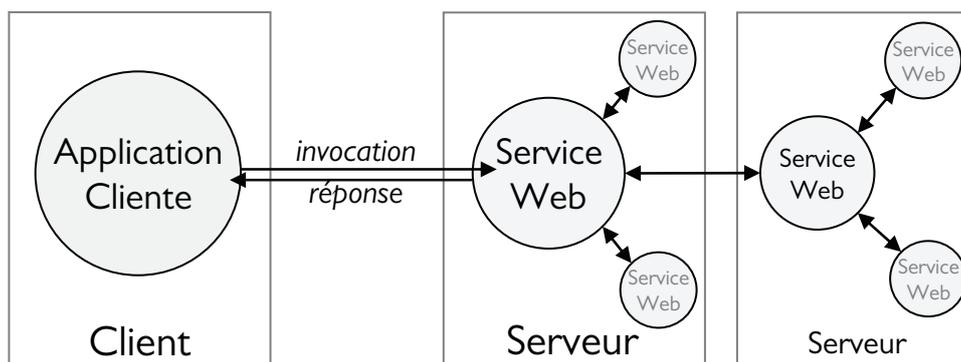


Figure 97 - Principe des services web.

¹ <http://www.w3.org>

4.2.3 L'évocation du service

Notre service de cartographie sémantique est évocable par la transmission d'un ordre de cartographie en MDL via le protocole SOAP.

a. Les ordres de cartographie

Dans sa version actuelle, Os Map n'accepte que des requêtes de cartographie décrite avec notre langage MDL (il est décrit en page 137). Ainsi, le service prend en entrée une demande de cartographie écrite en MDL et retourne l'URI de la carte construite. Comme décrite dans la section sur MDL, une cartographie peut contenir plusieurs cartes constituant autant de points de vue sur l'espace informationnel structuré.

L'opération d'invocation du service est nommée « map » et elle accepte les différents paramètres suivants :

- > Une **cartographie** : les demandes de cartographies sont décrites en MDL.
- > Un **nom** : pour nommer la carte.
- > Une **version** : pour permettre la gestion de différentes versions d'une même cartographie.
- > Un **titre** : le titre est automatiquement rajouté à la carte.
- > Un **auteur** : le système contrôle qui a accès (et avec quels droits) au service.
- > Une **description** : la carte peut posséder une description en langage naturel.
- > Une **date** : la réalisation d'une carte peut être programmée à une date donnée (utile si la carte nécessite d'importants traitements).
- > La **persistance** : il peut être demandé au système de conserver la carte pour permettre des utilisations ultérieures à la première utilisation.

Toutes les requêtes soumises au système de cartographie sont stockées pour permettre d'analyser les requêtes qui entraînent un éventuel dysfonctionnement du système mais aussi, pour effectuer ultérieurement des analyses statistiques sur l'usage du service.

b. La description du service

Os Map utilise WSDL comme langage de description de services. WSDL (« Web Service Definition Language ») est un langage de description de services web appartenant à la famille des langages XML [Gardarin, 2002; Kadima & Monfort, 2003]. Il a pour objectif de faciliter l'interopérabilité des services web. WSDL permet de décrire un service ainsi que les moyens pour y accéder. Plus particulièrement, WSDL permet de décrire l'ensemble des opérations invocables par les clients ainsi que la structure des différents messages. Il supporte un modèle d'échange inspiré de la navigation maritime : des messages contenus dans des containers (les enveloppes) sont échangés de port à port (port de communication du client et du serveur). À chaque opération sont associés deux messages : un premier pour invoquer le service et un second pour véhiculer le résultat.

La description complète du service en WSDL est disponible en annexe de ce document (voir à la page 248).

4.2.4 La mise en œuvre

En accord avec notre proposition d'architecture orientée service, Os Map a été construit comme un système multiagents. Il est composé des différents agents identifiés dans notre proposition d'architecture (en page 147) mais avec un agent supplémentaire : l'« atlas ». Ce dernier lui permet de gérer les cartes sur le serveur.

Les agents sont décrits dans les sections suivantes et ils sont répartis selon qu'ils interviennent sur le serveur ou sur le client.

a. Les agents côté serveur

Le serveur supporte une grande partie du service de cartographie sémantique. Il est en charge de supporter la construction et l'hébergement des cartes.

i. L'agent Service de Cartographie Sémantique

Os Map est l'agent responsable du service web. Il écoute le réseau en attente de requête de cartographie. À chaque nouvelle requête, il vérifie sa validité et il la stocke dans son historique de requêtes. Ensuite, il crée un agent Cartographe et le charge de réaliser la cartographie.

ii. L'agent Cartographe

L'agent Cartographe est responsable de la cartographie. Il est associé à une requête et un client uniques.

La construction de la carte se déroule en quatre étapes :

1. *Gestion des métadonnées* : le cartographe récupère les différentes métadonnées (comme le nom de la carte, son titre, son auteur et sa durée de validité) et il vérifie que l'utilisateur qui demande une cartographie a bien les droits nécessaires.
2. *Construction de l'espace informationnel structuré* : le cartographe délègue la construction de l'espace à cartographier à un agent Espace Informationnel Structuré (EIS).
3. *Cartographie de l'espace informationnel structuré* : pour rappel, une cartographie peut produire une ou plusieurs cartes. Le cartographe identifie alors les différentes sous-cartographies et pour chacune, il délègue l'application des différentes opérations de cartographies à un agent Carte.
4. *Publication de la carte* : le cartographe collecte les cartes (de chaque sous-cartographie) et les publie dans l'atlas. Ensuite, il les assemble en une seule carte selon la disposition décrite par la requête MDL. Cette carte globale se traduit par une page HTML contenant les différentes cartes. Elle est aussi publiée dans l'atlas et le cartographe retourne l'URI de la carte retourné par l'agent Atlas.

Dans toute requête MDL, à chaque carte est associée une sous-cartographie (la carte finale étant l'association de toutes les cartes). Les cartes sont visualisées sur le poste client. Par conséquent sur le serveur, il suffit de construire pour chaque carte l'espace informationnel représenté qui précède sa visualisation. Ainsi, l'agent Carte (placé sur le client) est chargé d'appliquer le reste des opérations de visualisation pour visualiser la carte construite « côté serveur ».

iii. L'agent EIS

L'agent EIS est chargé de construire l'espace informationnel *structuré* associé à la requête de cartographie. L'agent EIS utilise notre formalisme de description (voir à la page 109). Ainsi, les connaissances sont décrites par un ensemble de schémas (les entités à cartographier) et un ensemble de relations dont les tuples décrivent les relations entre les schémas.

iv. l'Atlas

Comme nous l'avons décrit précédemment, les cartes ne sont pas transmises en retour à une demande de cartographie. Os Map propose un service supplémentaire d'hébergement de cartes. Ainsi, à chaque cartographie, le service retourne un URI.

L'Atlas est l'agent responsable de l'hébergement des cartes. Pour le système de cartographie, il a deux fonctions : publier les cartes et les rendre accessibles.

La publication d'une carte est provoquée soit, par un agent Carte lorsqu'il a terminé d'appliquer les différentes opérations de cartographie soit, par un agent Cartographie lorsqu'il construit la carte globale contenant les différentes cartes. Toute demande de publication se traduit par l'envoi à notre agent Atlas d'une carte (un flux de caractères). Ce dernier stocke alors chaque carte dans un fichier selon son organisation interne et retourne l'URI associé. Techniquement, l'Atlas est basé sur un serveur Internet de fichiers classique et les ressources sont accessibles par le protocole HTTP.

L'accès à une carte par un utilisateur (généralement via son agent Carte sur le poste client) s'effectue avec l'URI de la carte comme avec tous les serveurs de fichiers.

Régulièrement, l'Atlas supprime les cartes inutiles. Une carte est jugée inutile en fonction de sa durée d'hébergement (paramétrée dans le système) et de sa durée de validité (soumise lors de la requête).

b. Les agents côté client

L'objectif d'Os Map est de cartographier un espace informationnel et de fournir des cartes accessibles et exploitables via Internet. Dans ce dessein, les cartes sont élaborées pour être visualisées sur le poste client (avec un navigateur par exemple).

Du côté client, le système dispose de deux types d'agents qui ont pour fonction de visualiser les cartes construites par Os Map : les visionneuses et le contrôleur.

i. Les visionneuses

Les visionneuses correspondent aux agents Cartes de notre proposition d'architecture et elles sont chargées de visualiser les cartes. Elles sont dédiées à chaque type de cartes, mais leur fonctionnement reste similaire. En effet, dans tous les cas il s'agit de pages HTML qui chargent une extension du système (comme flash, SVG, etc.) ainsi que la carte à visualiser. Le système (via l'extension concernée) se charge alors de visualiser la carte visible.

Dans la version actuelle d'Os Map, les principales visionneuses sont :

- X3D Viewer et SVG Viewer : ces visionneuses permettent de visualiser respectivement les cartes au format X3D et au format SVG.
- Les « Applet » Viewer : ces visionneuses permettent de visualiser toutes les cartes « métiers » c'est-à-dire les cartes qui nécessitent une application dédiée comme Os EyeTree ou Os RadialTree (voir à la page 189).

Chaque carte évolue indépendamment des autres. De plus, pouvant être des applications tierces, elles n'ont pas été conçues pour communiquer entre elles. La difficulté est alors de garantir la cohérence de l'ensemble.

ii. Le contrôleur

Le contrôleur correspond à l'agent Gestionnaire de Carte dans notre proposition d'architecture ; il permet d'assurer la cohérence entre les différentes cartes et leur espace informationnel. Toutes les interactions passent par le contrôleur. Il transmet alors chaque requête à l'agent Cartographe (sur le serveur). En fonction du résultat, il met à jour les cartes concernées.

Techniquement, à chaque fois qu'une carte souhaite interagir avec son espace informationnel, elle recharge la page associée au contrôleur en passant sa requête en paramètre. Ceci s'effectue comme une simple navigation hypertexte. Le contrôleur envoie la requête au serveur et attend la réponse. Des mécanismes de JavaScript lui permettent alors de mettre à jour les cartes si besoin.

4.2.5 Exemples

Cette section présente quelques exemples de cartes réalisées avec Os Map. Elles sont accessibles via un navigateur Internet.

La première carte est une cartographie en 3D d'un système de fichiers :

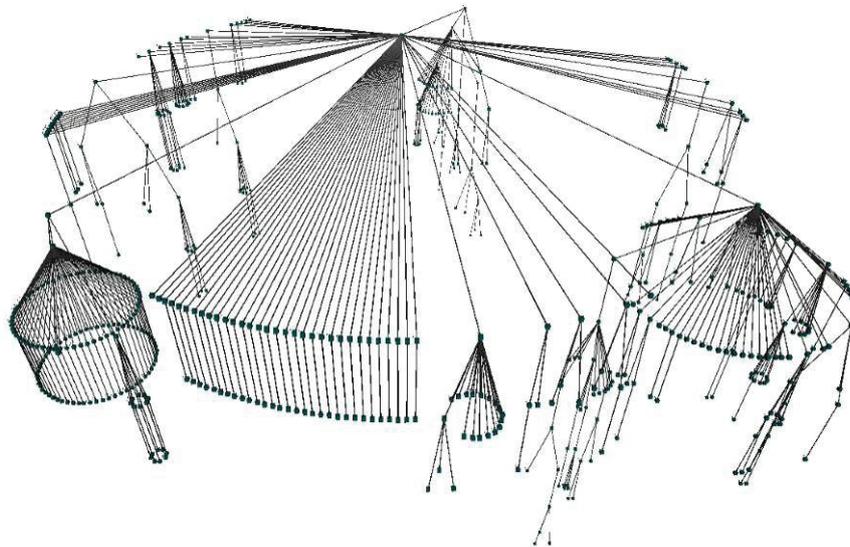


Figure 98 - Os Map : cartographie d'un système de fichiers.

Cette cartographie met en œuvre la technique des arbres de cônes (voir en page 68). Elle est en 3 dimensions et exploite le langage X3D¹. L'utilisateur peut déplacer l'arbre et naviguer dans tout l'espace de la carte.

La carte suivante est un autre exemple de carte réalisée avec Os Map. L'espace informationnel cartographié ici est « le référentiel des compétences des métiers des eaux et

¹ <http://www.web3d.org>

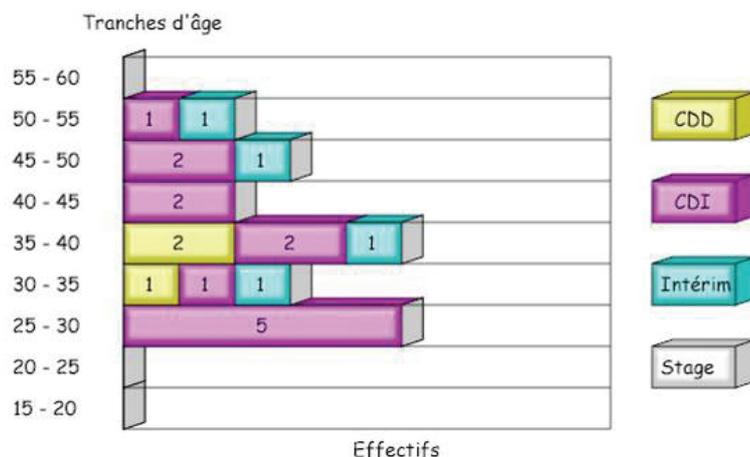


Figure 100 - Os Map : cartographie d'une pyramide des âges.

Cette cartographie utilise le langage SVG.

Tout comme les exemples précédents, l'utilisateur peut sélectionner les différents éléments de la carte pour avoir plus de détails (technique d'interaction de type « détails à la demande » ; voir à la page 87).

4.2.6 Synthèse

Os Map est notre réalisation d'un service web de cartographie sémantique. Il a été conçu pour être intégré à toute application qui souhaite cartographier des espaces informationnels. Il est la mise en œuvre directe de nos propositions. Tout d'abord, il exploite notre méthode et notre processus de cartographie sémantique. Ensuite, il implémente une version de notre formalisme et de notre architecture à agents. Enfin, les cartes qu'il permet de construire sont décrites avec MDL, notre langage de cartographie et elles sont visualisées entre autres avec nos deux paradigmes d'exploration, l'EyeTree et le RadialTree.

Os Map est donc un service de cartographie sémantique de haut niveau. Il a été conçu pour être intégré à toute application qui souhaite cartographier sémantiquement son espace informationnel.

4.3 Modélisation graphique avec SNCW

SNCW est un outil de cartographie sémantique pour modéliser graphiquement un domaine donné. Il exploite notre méthode de cartographie sémantique mais aussi notre formalisme de description et notre processus.

Toute notre approche de la gestion des connaissances repose sur la construction puis l'exploitation d'un référentiel métier par domaine de connaissances. Pour que le référentiel soit exploitable par la suite (dans d'autres applications par exemple), il est nécessaire qu'il repose sur une modélisation formelle. Dans notre approche, ce référentiel est une modélisation ontologique du domaine. Ce référentiel est construit selon la méthode Ousia et avec les outils du monde KCW (voir « La démarche Ousia » page 160).

Ce modèle théorique permet de structurer les connaissances du domaine et plus particulièrement ses concepts en une ontologie (modélisation formelle des connaissances).

L'enjeu de cette approche de gestion des connaissances est la construction de ce référentiel. Elle nécessite l'intervention des experts du domaine pour décrire les concepts du domaine. Les retours d'expérience de notre équipe dans ce domaine nous ont permis de comprendre qu'il est difficile de modéliser les connaissances d'un domaine. Cette constatation est d'autant plus vraie si la modélisation repose sur une théorie qui possède des contraintes fortes.

Dans notre cas, les référentiels métiers reposent sur des connaissances ontologiques et terminologiques décrites formellement (théorie sous-jacente fortement contrainte). C'est pourquoi, une étape préliminaire dans la construction du référentiel est nécessaire. Cette étape doit permettre aux experts de décrire leurs connaissances selon une théorie « semi-formelle » située entre l'expression « informelle » en langage naturel et la modélisation « formelle » à l'aide d'ontologies.

Cette étape préliminaire doit être supportée par un outil de modélisation « semi-formelle » permettant de modéliser toutes les connaissances du domaine et non uniquement les connaissances qui vont appartenir au référentiel formel (des connaissances ontologiques et terminologiques dans notre cas).

Face à cette problématique de modélisation d'un domaine, nous avons souhaité exploiter notre approche de la cartographie sémantique pour définir un nouvel environnement de modélisation graphique des connaissances. Nous l'avons nommé SNCW pour « Semantic Network Craft Workbench ». Il sert d'outil intermédiaire entre LCW (Linguistics Craft Workbench) pour l'analyse linguistique des vocabulaires métier et les environnements OCW (Ontology Craft Workbench) et TCW (Terminology Craft Workbench) pour leur définition formelle (voir à la page 160).

SNCW permet aux experts de modéliser leur domaine avec pour finalité la construction d'un système d'entités structurées capable d'exprimer leurs connaissances. Ce système produit alors ce que nous nommons une « base de connaissances ». SNCW est notre environnement graphique et interactif pour la modélisation et la représentation semi-formelle de connaissances.

4.3.1 La modélisation graphique de connaissances

Certains environnements de modélisation ont recours à des techniques graphiques pour cartographier leurs bases de connaissances.

Dans la plupart des cas, ces environnements utilisent la cartographie comme un mécanisme permettant de modéliser au travers de représentations graphiques des concepts théoriques. C'est par exemple le cas avec les environnements de modélisation logicielle qui proposent de construire graphiquement des diagrammes de classes. Les classes et leurs associations, constituant les principaux concepts théoriques des diagrammes de classes, sont représentées respectivement par des rectangles et par des traits.

Avec une représentation graphique des concepts théoriques, l'utilisateur de ces environnements peut aisément manipuler ces concepts tout en assimilant inconsciemment leur sémantique (sous réserve du choix et de la qualité des représentations utilisées). L'utilisateur n'a alors plus qu'à se concentrer sur la construction de sa modélisation en elle-même.

Les environnements de modélisations graphiques peuvent être répartis en deux classes selon qu'il propose une théorie dite « formelle » ou « semi-formelle ». Les sections suivantes présentent quelques-uns de ces environnements.

a. Les approches par modélisation formelle

Les environnements de modélisation formelle qui proposent une cartographie sont relativement peu nombreux. Néanmoins, ce constat tend à changer comme en témoigne l'environnement « Protégé¹ » qui propose depuis peu des représentations graphiques.

Protégé est un éditeur de bases de connaissances, dites ontologiques, offrant deux approches différentes avec deux théories supportées. La première approche utilise une théorie orientée systèmes à base de schémas et la seconde, utilise une théorie issue d'OWL². L'environnement supporte l'ajout d'extensions pour visualiser les modélisations. Toutes ces extensions exploitent des représentations de type nœud-lien pour représenter les concepts théoriques. Si nous prenons l'approche OWL, les nœuds représentent des classes ou des instances et les traits représentent les relations entre les classes et les instances.

Voici un exemple de représentation proposé par l'extension OWLViz³ pour Protégé :

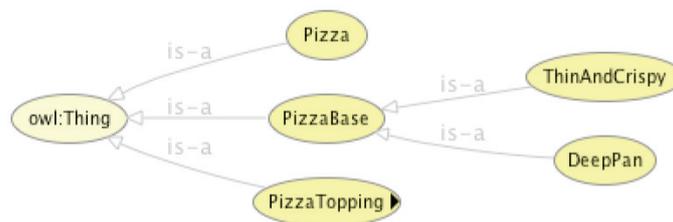


Figure 101 - Modélisation graphique avec Protégé.

¹ <http://protege.stanford.edu>

² OWL est un langage de description d'ontologies pour le web sémantique supporté par le W3C (<http://www.w3.org/2001/sw/WebOnt>).

³ <http://www.co-ode.org/downloads/owlviz>

De même, OCW (voir à la page 161) a recours à des cartographies pour construire des modélisations. Ainsi, il permet de construire des ontologies basées sur un système de concepts liés par une relation de généralisation/spécialisation.

Voici un exemple de cartographie d'ontologie réalisée par OCW :

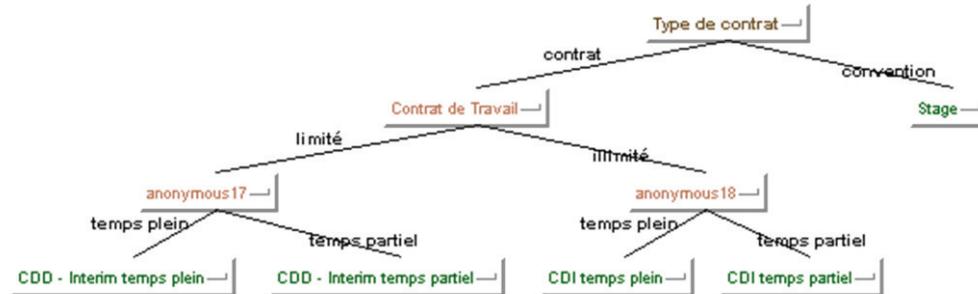


Figure 102 - OCW : cartographie d'ontologie.

Dans ces deux exemples, la carte contient une représentation de l'ensemble des concepts (ou classes) de la modélisation. Étant donné que la nature de la théorie utilisée est « formelle », l'utilisateur ne peut que rajouter de nouveaux concepts (ou classes) et les lier entre eux par la seule relation autorisée. De plus, le système effectue de nombreux contrôles de cohérence qui rendent la création d'un concept plus complexe que le simple ajout d'une représentation. On comprend alors la difficulté dans une première démarche de modélisation de se limiter à un cadre aussi contraint.

De plus, dans ces outils de modélisation formelle, les concepts manipulés pour décrire ne permettent pas à l'utilisateur d'exprimer des connaissances de nature différentes. Cependant, dans la démarche de construction, il est très utile pour commencer par « tout poser à plat » et ensuite de structurer les informations. Ce constat nous guide donc vers des outils de modélisation qualifiés de « semi-formels ».

b. Les approches par modélisations semi-formelles

Les solutions de modélisations avec une théorie semi-formelle sont beaucoup plus nombreuses. Les plus courantes sont les outils pour la construction de cartes heuristiques (« Mind Mapping ») et ceux pour la construction de cartes conceptuelles (« Concept Mapping ») [Neumann et al., 2005].

L'objectif des cartes heuristiques est d'aider les utilisateurs à développer graphiquement une idée principale en un ensemble de sous-idées. L'idée principale correspond au thème de la carte. Elle est placée au centre de la carte et les idées qui en dérivent sont placées de part et d'autre. Les concepts théoriques associés à cette carte sont limités à un ensemble d'entités (correspondant aux idées) et à des liens entre ces entités. Aucune sémantique n'est associée formellement à ces liens.

Voici un exemple de carte réalisée avec FreeMind¹ :

¹ <http://freemind.sourceforge.net>

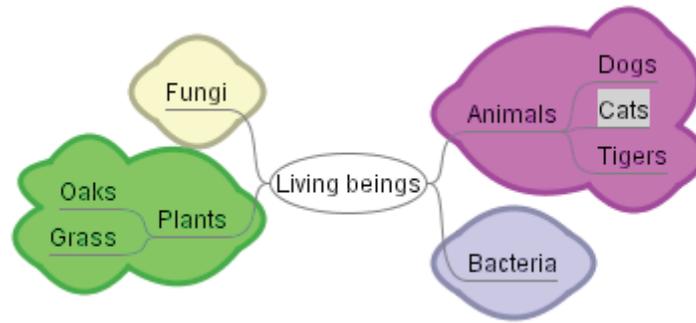
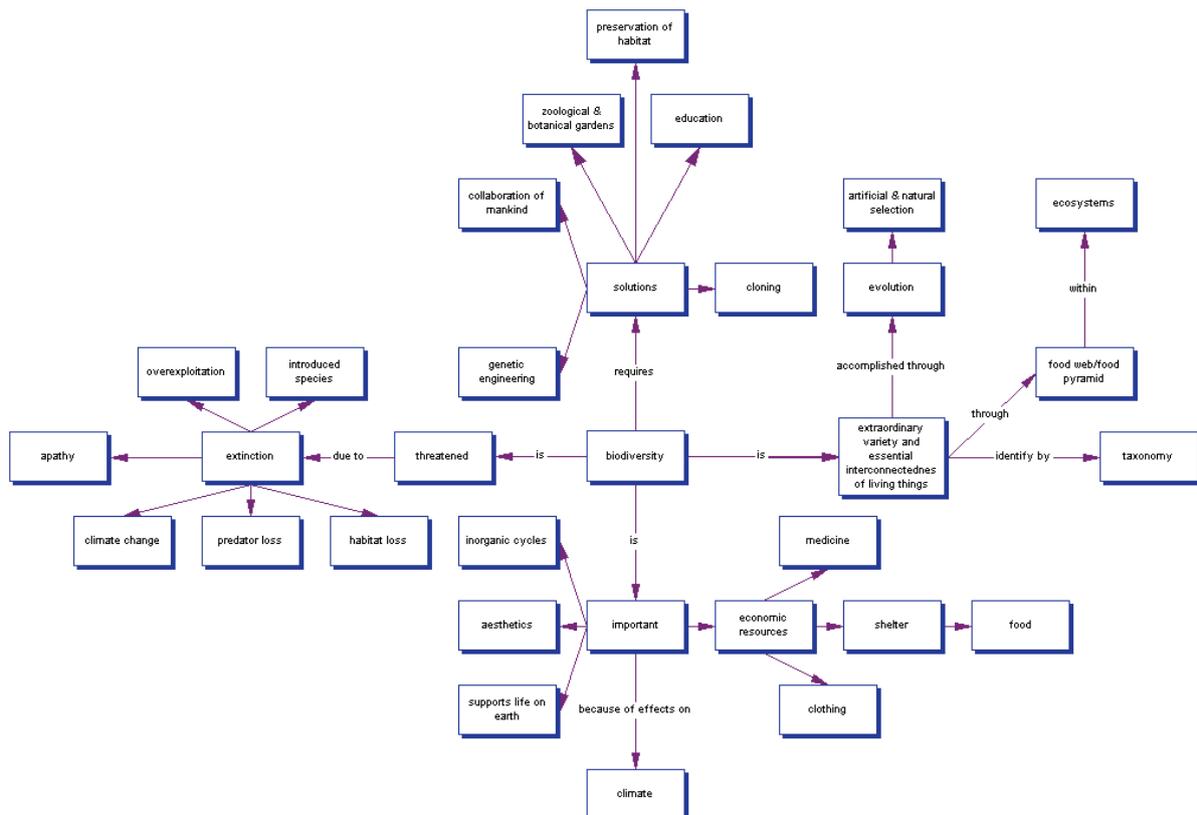


Figure 103 - Exemple de carte heuristique.

Les cartes heuristiques sont des représentations d'arborescences d'idées généralement utilisées pour prendre des notes, préparer des exposés, raisonner à plusieurs sur un thème (brainstorming) ou bien encore, pour l'apprentissage.

Les cartes conceptuelles quant à elles, sont généralement utilisées pour représenter graphiquement la modélisation d'un domaine telle qu'elle est perçue par un ou plusieurs individus. Une carte conceptuelle contient un ensemble d'idées nommées « concepts » représentées par des nœuds et ils sont associés les uns aux autres par des relations (éventuellement prédéfinies).

Voici un exemple de carte conceptuelle :

Figure 104 - Exemple de carte conceptuelle sur la biodiversité¹.

¹ <http://www.woodrow.org/teachers/bi/1999/projects/group4/Roberts>

Tout comme les cartes heuristiques, les cartes conceptuelles sont composées d'un ensemble de nœuds (forme circulaire ou rectangulaire) représentant les concepts et un ensemble de traits représentant les relations entre les concepts. Généralement, les outils de construction de cartes conceptuelles permettent d'utiliser des relations prédéfinies ou d'en définir de nouvelles. À remarquer qu'aucune sémantique n'est associée aux différentes relations (uniquement une étiquette et éventuellement une représentation différente).

Dans la même classe des cartographies de modélisations semi-formelles, on peut noter que certaines extensions associées aux outils de « Topic Maps¹ » permettent d'obtenir une cartographie relativement proche des cartes conceptuelles.

La présentation de ces trois familles de modélisations semi-formelles permet de révéler leurs limites par rapport à nos besoins : construire une modélisation servant d'étape préliminaire à la construction d'un référentiel métier (modélisation formelle).

Nous pouvons classer ces limites sur deux plans : théorique et graphique.

Sur le plan théorique, les approches de modélisation semi-formelle ne permettent pas de garantir un minimum de cohérence pour les modélisations construites. La cohérence ne peut être contrôlée que si les éléments de la modélisation possèdent une sémantique. Par exemple, il n'est généralement pas autorisé de définir une nature pour les éléments décrits (concepts, idées ou entités) et par conséquent, il n'est pas possible de contrôler la cohérence lors de la création de liens entre ces éléments.

Dans une démarche personnelle de réflexion, le contrôle de la cohérence n'est pas indispensable. Cependant, nous souhaitons fournir un outil capable d'aider les utilisateurs à définir une modélisation réutilisable pour construire un référentiel métier. Un contrôle de la cohérence de la modélisation est alors indispensable.

Sur le plan graphique, ces outils de modélisation proposent de construire des cartes avec une seule vue. Une vue est largement suffisante pour de petites modélisations mais avec des cas issus du monde des entreprises, les cartes sont vite surchargées et l'utilisateur ne peut plus appréhender l'ensemble de la modélisation. Il est alors nécessaire de proposer une solution pour ne visualiser qu'une partie de la modélisation en fonction de certains critères comme la nature des entités ou le type de relations qui les unissent.

4.3.2 Besoins pour la modélisation graphique

Notre objectif est de fournir un environnement de modélisation graphique pour la construction de bases de connaissances. Les sections suivantes présentent tout d'abord les besoins pour spécifier un environnement de modélisation graphique.

a. Plusieurs modélisations d'un même monde

Comme nous l'avons décrit dans notre processus (voir à la page 101), pour modéliser un domaine de connaissances (tacites ou explicites) il est souvent nécessaire de pouvoir gérer plusieurs modélisations simultanément. Chaque collaborateur possède sa propre vision de son domaine, par conséquent chaque collaborateur possède aussi ses propres connaissances sur ce domaine.

¹ <http://www.topicmaps.org>

Pour illustrer la possibilité de concevoir plusieurs modélisations d'un même monde, prenons l'exemple de la lumière. Dans un premier cas, la lumière peut être modélisée par un modèle corpusculaire : « la lumière est un flux qui contient des particules ». Dans un second cas, la lumière peut être modélisée par un modèle ondulatoire : « la lumière est une onde ». Ces deux modélisations produisent donc bien deux modèles différents.

Dans une démarche préliminaire à la définition du référentiel, il est nécessaire de pouvoir confronter ces différentes modélisations. Notre environnement doit donc être capable de supporter **plusieurs modélisations d'un même « monde »**, le monde étant ici constitué des connaissances de l'organisation (tacites et explicites).

Les différentes modélisations d'un même monde sont les « points de vue théoriques ».

b. Plusieurs vues d'une même modélisation

Une modélisation peut contenir un très grand nombre d'informations, d'autant plus si elle ne se limite pas à certaines natures de connaissances. L'utilisateur (même expert du domaine) ne peut pas facilement appréhender l'ensemble des données dans sa globalité (même avec le recours à la visualisation). Il est alors nécessaire de proposer plusieurs vues d'une même modélisation pour que les utilisateurs puissent appréhender chaque modélisation soit dans sa globalité soit de manière plus ciblée.

Pour illustrer ce besoin d'avoir plusieurs vues d'une même modélisation, prenons l'exemple du domaine de la cartographie de données abstraites. Imaginons un expert qui souhaite modéliser ce domaine en structurant tous ses concepts. On imagine vite le nombre de concepts à représenter. Étant donné que la cartographie met en jeu plusieurs domaines de compétences, une approche pourrait consister à cartographier chaque domaine dans une vue dédiée ainsi que tous les concepts qui s'y rapportent. Les experts de chaque domaine pourraient alors se concentrer sur leur domaine. Une vue pourrait aussi être utilisée pour visualiser toutes les entités du modèle liées par une relation particulière.

L'utilisation de plusieurs vues pour une même modélisation présente de nombreux avantages. Il nous semble donc nécessaire de proposer un environnement capable de supporter **plusieurs vues d'une même modélisation**.

Notre principe de modélisation graphique correspond donc à ce schéma :

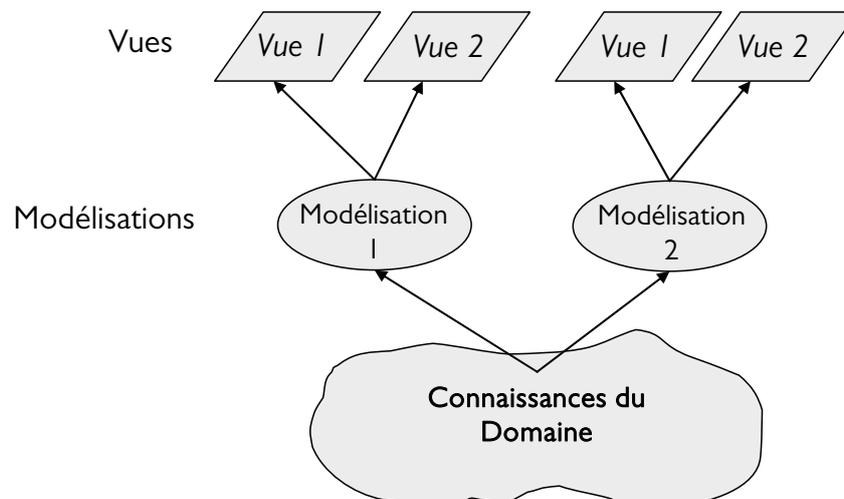


Figure 105 - SNCW : besoins pour la modélisation graphique.

La cartographie sémantique permet de répondre à ces besoins.

4.3.3 Le principe de SNCW

Notre objectif est de fournir un environnement de modélisation graphique pour la construction de bases de connaissances.

Notre processus de cartographie sémantique permet de modéliser un monde par plusieurs modélisations simultanément. La particularité dans ce cas est que l'espace informationnel n'est pas structuré sur cette modélisation mais c'est la modélisation qui est cartographiée. La modélisation est alors assimilable à l'espace informationnel *structuré* de notre processus. Notre processus nous permet d'avoir une multitude de cartes de l'espace à cartographier, par conséquent notre processus permet d'avoir plusieurs vues de la modélisation.

Nous savons aussi que notre environnement doit décrire plusieurs modélisations d'un même monde mais aussi, plusieurs vues d'une même modélisation. La théorie que nous avons retenue pour notre environnement est donc celle que nous proposons avec notre formalisme. Ainsi, les connaissances sont exprimées par un ensemble de modélisations décrites par des réseaux sémantiques (un ensemble d'entités et de relations dont les tuples mettent les entités en relations).

Reste alors à déterminer le choix des paradigmes pour cartographier ces réseaux sémantiques correspondant aux différentes modélisations afin d'obtenir les différentes vues. Nous savons qu'une vue est la visualisation d'un espace *représenté* contenant des structures visuelles. D'après notre approche de la cartographie sémantique, ces structures visuelles doivent respecter la sémantique de la modélisation. Étant donné que chaque modélisation est un réseau sémantique, les représentations de type nœud-lien sont préconisées.

Un réseau sémantique est constitué d'entités et de relations entre ces entités. Par conséquent, les vues doivent contenir des structures visuelles pour les entités et pour les tuples (les relations sont représentées au travers de leurs tuples). D'un point de vue descriptif, une vue est constituée de deux ensembles : les représentations d'entités et les représentations de tuples. Les représentations de tuples mettent visuellement en relations les représentations d'entités.

Face à ce constat, la similarité avec la description des modélisations nous a paru évidente. Par conséquent, nous avons décidé d'utiliser aussi notre formalisme pour décrire l'espace *représenté* des modélisations (correspondant aux vues).

On obtient donc le schéma suivant pour le principe de SNCW :

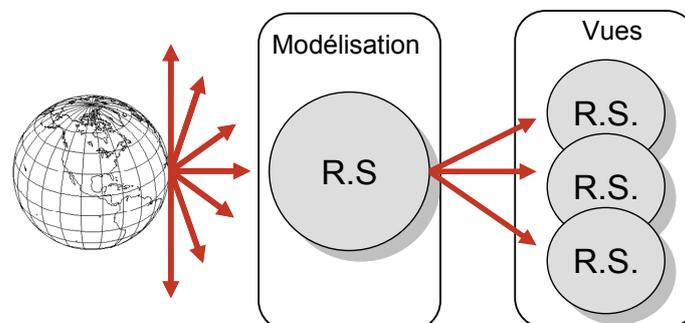


Figure 106 - Principe de SNCW.

Les bases de connaissances construites avec SNCW sont composées de modélisations et de cartographies de ces modélisations, les vues ; les modélisations et leurs représentations sont décrites par des réseaux sémantiques.

a. Modélisation

Avec SNDF, une modélisation est une description semi-formelle de connaissances, c'est-à-dire des objets du domaine et des concepts qui les décrivent. Chaque modélisation correspond à « un point de vue théorique » des objets du domaine. La description de ces connaissances est donnée sous la forme d'un réseau sémantique, c'est-à-dire sous la forme d'un ensemble d'entités liées par des relations (arcs orientés étiquetés).

Voici un exemple de modélisation réalisée avec notre environnement SNCW :

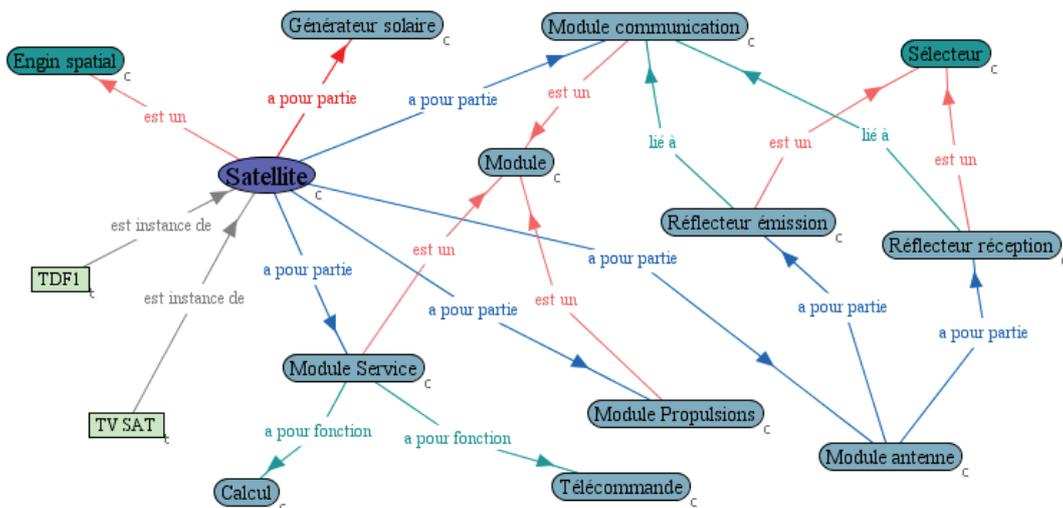


Figure 107 - SNCW : modélisation de satellite.

b. Vue

Chaque vue est la visualisation de la représentation graphique d'une modélisation (du réseau sémantique). Pour reprendre la terminologie associée à notre processus, chaque vue est la modélisation *représentée* puis la modélisation *visualisée*. Lorsqu'un réseau est complexe, en nombre d'entités et de relations, il est utile de ne visualiser que certaines entités ou certaines relations. C'est pourquoi SNCW permet de gérer plusieurs vues pour une même modélisation où chaque vue contient un sous-ensemble de la modélisation.

C'est le cas par exemple avec la saisie d'écran suivante :



Figure 109 - SNCW : attributs de représentation.

Les éléments de la modélisation peuvent avoir plusieurs représentations dans des vues différentes. SNCW offre alors un mécanisme simple pour créer une nouvelle représentation d'une entité, d'un ensemble d'entités ou de tuples dans une autre vue :



Figure 110 - SNCW : création de représentations.

c. Base de connaissances

SNCW permet de créer des bases de connaissances (décrites par SNDF). Nous nommons « base de connaissances » toutes les modélisations d'un même domaine d'application.

Dans ce contexte, chaque modélisation permet de modéliser un « point de vue théorique » sur un domaine. Ainsi un même « objet » (instance), s'il est lié à un seul concept par la relation « instance » dans une modélisation donnée (ou du moins devrait l'être), peut être lié à un autre concept par la relation « instance » dans une autre modélisation. Le même objet peut être appréhendé (conceptualisé) selon plusieurs modélisations différentes. En revanche, un même concept ne peut appartenir à plusieurs modélisations, il appartient à une modélisation donnée et à une seule.

Notre formalisme permet à SNCW d'assurer cette cohérence en permettant de partager simplement (et uniquement) les objets entre les modélisations.

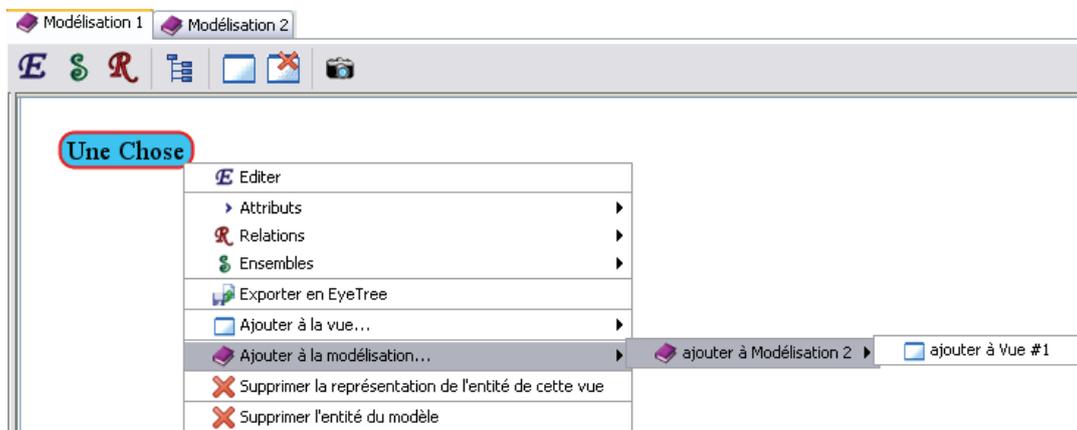


Figure 111 - SNCW : partage d'objets entre les modélisations.

Tout comme le partage d'éléments entre les différentes vues, SNCW permet de partager les objets entre les modélisations.

4.3.4 L'architecture de SNCW

Pour mettre en œuvre notre environnement de modélisation, nous l'avons scindé en deux parties. La première correspond au module de gestion des bases de connaissances (« SNCW KB Manager ») et la seconde (« SNCW Editor ») correspond à l'interface entre l'homme et le module de gestion de bases.

Remarque : ce choix d'organisation permet d'anticiper la probable évolution de SNCW vers un service web de modélisation graphique.

Pour le module de gestion de bases, nous avons opté pour une architecture à agents logiciels similaire à notre architecture (voir à la page 147). Pour communiquer entre eux, les agents utilisent deux langages : LSNCW et LSN. Ils sont décrits dans les sections suivantes.

Le schéma suivant présente notre architecture en deux parties et les différents agents qui la composent :

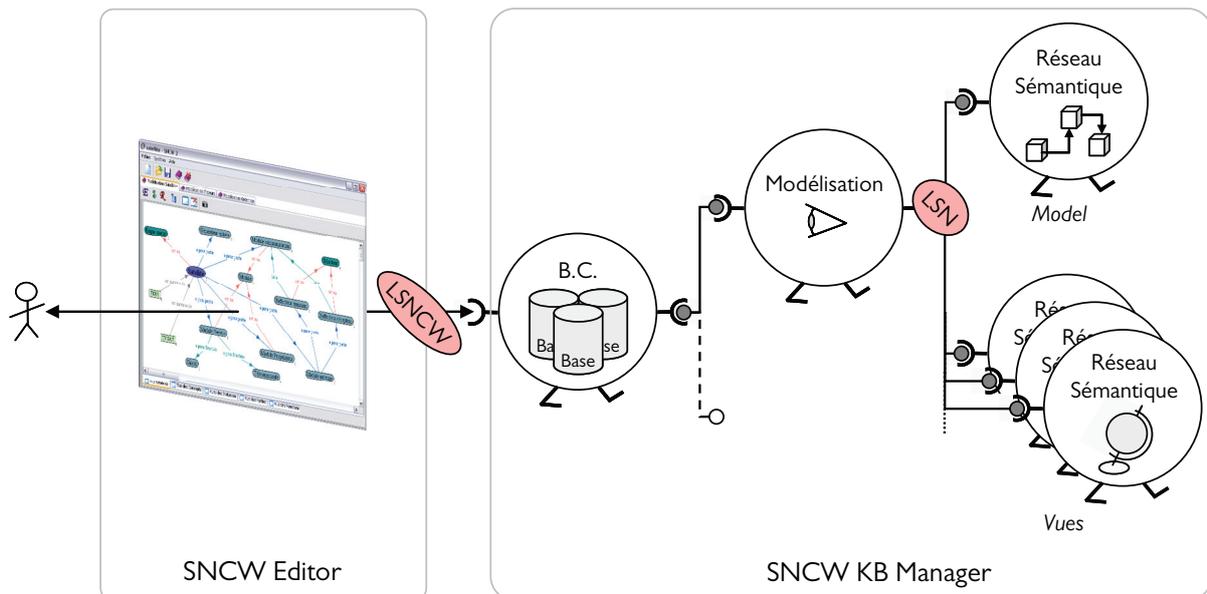


Figure 112 - Architecture de SNCW.

Dans SNCW, le réseau sémantique qui décrit le contenu de la modélisation est nommé « model ».

a. Les agents de base de connaissances

À chaque base de connaissance est associé un agent dont le rôle est de gérer les différentes modélisations qui composent la base. Cet agent est aussi en charge de différents paramètres communs à l'ensemble des modélisations comme les paramètres d'édition de l'utilisateur (forme par défaut, couleur par défaut, etc.).

De plus, les agents de base disposent d'un historique des opérations effectuées sur les bases pour permettre des opérations d'annulation ou de répétitions.

b. Les agents de modélisation

À chaque modélisation est associé un agent de modélisation. Ce dernier a alors pour fonction de gérer le réseau sémantique qui décrit le contenu de la modélisation (nommé « model ») et il doit aussi gérer les réseaux sémantiques qui correspondent aux vues.

Il permet aussi d'assurer la communication entre l'agent de base de connaissances et les réseaux. À ce titre, il communique avec les différents agents de réseau sémantique en LSN (décrit dans la section suivante).

c. Les agents de réseau sémantique

Chaque agent de réseau sémantique est en charge de gérer un réseau qui permet soit décrire une modélisation (nommée « model ») soit de décrire une vue (représentation de modélisation). Ils communiquent avec leur agent de modélisation au moyen du langage LSN (pour « Language for Semantic Network »).

Cette section a pour objet de présenter les fonctionnalités des agents de réseau sémantique au travers de la description des différents messages qui permettent de les invoquer. Tous ces messages composent LSN.

LSN est un langage de type XML. Pour des raisons de lisibilité, nous avons choisi de le présenter avec une syntaxe simplifiée : les balises sont remplacées par des parenthèses et les paramètres sont indiqués entre côtes simples. La flèche (« → ») fournit des précisions sur le résultat après invocation du service associé. Cette présentation du langage n'est pas aussi rigoureuse que le langage qu'elle présente mais elle a pour objectif d'illustrer simplement sa description.

La première action possible correspond à la création du réseau sémantique. Chaque réseau possède un nom. La grammaire du message qui permet de l'invoquer est donc la suivante :

```
(defineSemanticNetwork 'name') → 'name'
Init: E = ∅, R = ∅, S={E}
```

À chaque création d'un réseau, celui-ci est nécessairement composé des ensembles E (pour les entités), R (pour les relations) et S (l'ensemble de tous les ensembles excepté R ; par conséquent il contient E).

i. Les entités

Les entités sont définies avec le message « defineEntity ». Chaque entité possède un nom unique qui permet de l'identifier. Toute entité créée appartient à l'ensemble E, l'ensemble de toutes les entités. Il est aussi possible de supprimer une entité à l'aide du message « deleteEntity ».

```
(defineEntity 'name') → 'name' ∈ E
(deleteEntity 'name')
```

Il est possible d'associer des attributs aux entités (comme dans les réseaux de schémas) au moyen du message « addSlot » et « addValue » en précisant le nom de l'attribut (slot) ainsi qu'éventuellement sa valeur. Le message « deleteSlot » permet de supprimer un attribut et le message « getValue » permet de récupérer sa valeur.

```
(addSlot 'entity' 'slot')
(deleteSlot 'entity' 'slot')
(addValue 'entity' 'slot' 'value')
(getValue 'entity' 'slot')
```

ii. Les ensembles

Les ensembles contiennent des éléments du réseau. Nous considérons les éléments comme étant soit des entités soit des relations soit des ensembles. Chaque ensemble (en dehors de S) appartient à S.

La création d'un ensemble se fait avec le message « defineSet » et l'ajout d'un élément se fait avec le message « addElement » et réciproquement pour les opérations inverses « deleteSet » et « deleteElementFromSet ».

```
(defineSet 'name') → 'name' ∈ S
(deleteSet 'name')
(addElement 'set' 'element') 'element' ∈ E ou ∈ R (element≠S)
(deleteElementFromSet 'set' 'element')
```

iii. Les relations

Les relations permettent de lier les entités entre elles. Toutes les relations appartiennent à l'ensemble « R ». Le message « defineRelation » permet de définir une nouvelle relation et réciproquement, le message « deleteRelation » permet de supprimer une relation.

```
(defineRelation 'name') → 'name' ∈ R
(deleteRelation 'name')
```

Toutes les relations possèdent un domaine défini par un vecteur d'ensembles.

```
(defineDomain 'relation' 's1'...'sk') 'sj' ∈ S avec 1<j<k
```

La création d'un tuple se fait avec le message « addTuple » en précisant les entités constituant le tuple. Les tuples de la relation doivent appartenir au produit cartésien des ensembles du domaine de la relation. Il est aussi possible de supprimer un tuple avec le message « deleteTuple ».

```
(addTuple 'relation' 'e1'...'ek') 'relation' ∈ R et 'ej' ∈ sj avec 1<j<k
(deleteTuple 'relation' 'e1'...'ek')
```

iv. Les macros

Les macros sont des messages de haut niveau qui permettent d'automatiser certaines tâches.

La première macro permet de définir un réseau sémantique pour décrire le contenu de la modélisation au moyen du message « defineNewModel ». Dans ce cas, un réseau sémantique est tout d'abord créé puis, deux ensembles (« Concepts » et « Things ») et trois relations (« est un », « est instance de » et « a pour partie ») avec leur domaine associé.

Remarque : les lignes indentées présentent la succession des opérations déclenchées par l'appel de la macro.

```
(defineNewModel 'myModel')
  (defineSementicNetwork 'myModel')

  (defineSet 'Concepts')
  (defineSet 'Things')

  (defineRelation 'est un')
  (defineDomain 'est un' 'Concepts' 'Concepts')
  (defineRelation 'est instance de')
  (defineDomain 'est instance de' 'Things' 'Concepts')

  (defineRelation 'a pour partie')
  (defineDomain 'est instance de' 'Concepts' 'Concepts')
```

La macro `defineNewView` permet de créer une vue en définissant tout d'abord un nouveau réseau sémantique puis deux ensembles (« Nodes » et « Tuples »).

```
(defineNewView 'myView')
  (defineSementicNetwork 'myView')
  (defineSet 'Nodes')
  (defineSet 'Tuples')
```

d. L'agent de gestion de bases de connaissances

L'agent de gestion de bases de connaissances nommé « KB Manager » est l'agent principal du système. Il constitue l'interface entre l'éditeur et les bases de connaissances. À ce titre, il est responsable de la communication avec les agents de base de connaissances.

Dans ce dessein, nous avons défini un langage de communication pour les agents de bases de connaissances et l'éditeur. Il constitue le langage de plus haut niveau dans notre environnement et nous l'avons nommé LSNCW (pour « Language for Semantic Network Craft Workbench »).

La fonction du langage LSNCW est donc de permettre à l'agent de gestion de bases de connaissances de recevoir des requêtes de l'éditeur pour manipuler des bases de connaissances. Dans ce nouveau langage, il est possible de distinguer deux types de messages : les messages qui portent sur la manipulation des bases en tant que bases (gestion externe) et les messages qui permettent de manipuler les éléments de ces bases (gestion interne).

i. Gestion externe des bases de connaissances

LSNCW offre quelques messages basiques pour manipuler les bases de connaissances. Tout d'abord, le message « `createNewKB` » permet de définir une nouvelle base de connaissances. Pour cela, il crée une modélisation par défaut nommée « ma modélisation » et pour cette modélisation, il crée un réseau sémantique pour décrire son contenu et un autre réseau pour une vue.

De plus, LSNCW offre les messages « `getKB` » et « `setKB` » qui permettent, d'une part de récupérer le contenu complet d'une base de connaissances pour la sauvegarder et d'autre part de soumettre un contenu pour la restaurer (le contenu est exprimé en XML).

Et enfin, LSNCW propose deux messages pour récupérer ou préciser des propriétés (couple nom et valeur) communes à toutes les modélisations d'une même base de connaissances comme les différents paramètres d'édition (par exemple forme et couleur par défaut).

ii. Gestion interne des bases de connaissances

Pour permettre à l'éditeur de manipuler les différents éléments des différentes bases de connaissances, LSNCW offre deux types de mécanismes : l'imbrication de commandes en LSN et des messages de plus haut niveau.

Imbrication de messages LSN

Le mécanisme d'imbrication permet d'imbriquer des messages LSN dans un message en LSNCW :

Lsn(messageLsn)

Ce mécanisme permet alors à l'agent éditeur d'appliquer les opérations proposées par LSN depuis l'interface sans recourir à la définition de nouveaux messages LSNCW pour y parvenir.

La mise en œuvre de ce mécanisme s'effectue par la transmission des messages en LSN à l'agent réseau sémantique concerné.

Messages de hauts niveaux

LSNCW définit de nombreux messages de hauts niveaux. Nous les qualifions ainsi car ils permettent de définir des opérations comme un ensemble d'opérations en LSN mais aussi en LSNCW. Ces messages de hauts niveaux utilisent donc le mécanisme d'imbrication de messages LSN.

Parmi ces messages de haut niveau, LSNCW propose des messages pour gérer les modélisations (comme l'ajout et la suppression de modélisations) et des messages pour gérer les réseaux sémantiques (comme la définition d'une entité, d'un ensemble ou la définition d'une vue). La définition de ces messages est relativement simple étant donné qu'ils ne font appel qu'à des messages en LSN (en précisant le réseau sémantique concerné).

Ensuite, LSNCW propose des messages de hauts niveaux plus complexes comme la suppression ou le changement de nom d'une entité, ou bien encore comme la suppression d'un ensemble.

Prenons par exemple le premier cas celui de la suppression d'une entité. Cette opération qui semble simple implique de nombreuses modifications « en cascade » de la modélisation concernée.

Voici la liste succincte des opérations engendrées par la suppression d'une entité « e » :

- > suppression de tous les tuples mettant en relation « e » avec une autre entité
 - > suppression de toutes les représentations des tuples concernées dans les différentes vues
 - > suppression de tous les tuples de la modélisation
- > suppression de toutes les représentations de l'entité « e » dans les différentes vues
- > suppression de l'entité « e » du modèle

Ainsi, LSNCW propose tout un ensemble d'opérations de haut niveau pour rendre transparente la gestion des bases de connaissances ; tout cela dans le but de permettre à l'agent éditeur de gérer simplement toutes les bases de connaissances.

e. L'agent éditeur

L'agent éditeur (« SNCW Editor ») joue le rôle d'interface entre l'utilisateur et l'agent gestionnaire de bases de connaissances (« SNCW KB Manager »). À ce titre, il est principalement constitué d'une interface graphique et d'un moteur capable de communiquer en langage LSNCW avec l'agent gestionnaire de bases.

i. L'interface graphique

L'interface de SNCW a été construite à l'image des bases de connaissances qu'il permet de construire. Ainsi, l'interface principale est composée d'un ensemble d'onglets à raison d'un onglet par modélisation. Les actions communes aux modélisations sont accessibles via une barre d'outils placée juste au-dessus de ces onglets.

De même, chaque onglet de modélisation est composé d'un ensemble d'onglets correspondant à chaque vue et d'une barre d'outils donnant accès à l'ensemble des opérations propres à la modélisation courante. De plus, chaque modélisation offre la possibilité de parcourir son contenu selon une vue arborescente classique (parcours de la modélisation indépendamment de ses vues).

L'éditeur propose l'interface suivante :

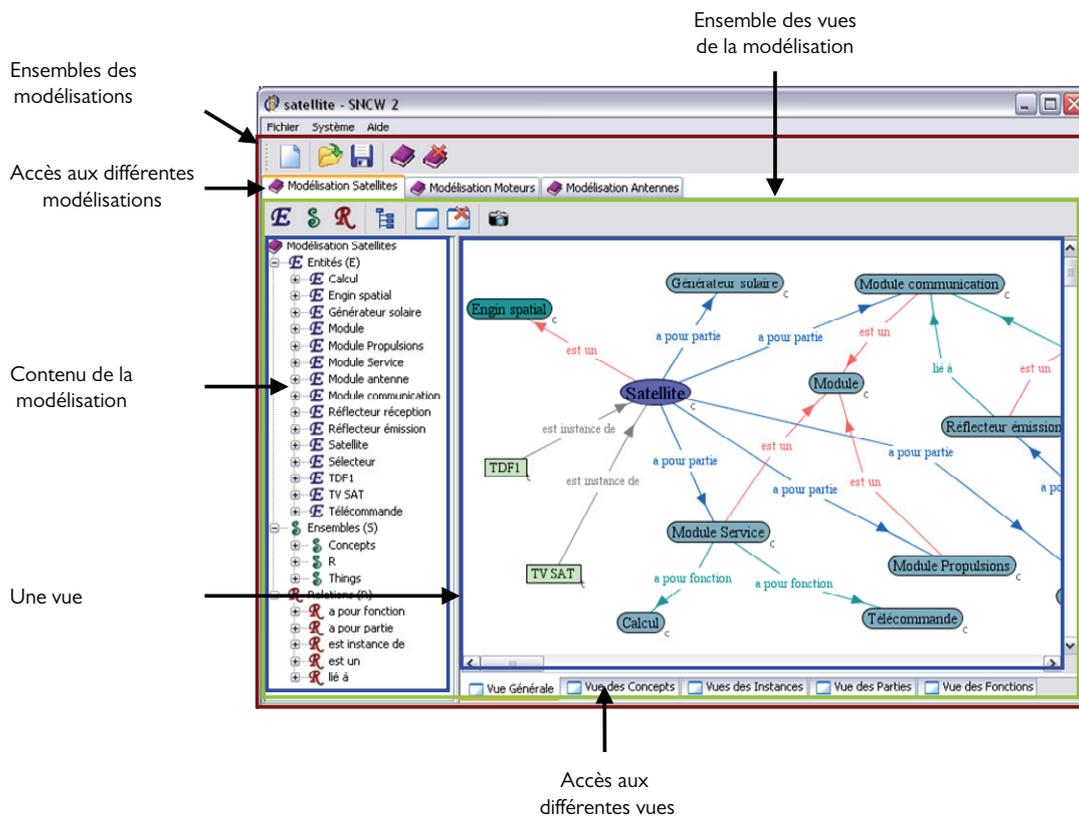


Figure 113 - Interface de SNCW.

ii. Interactions avec l'agent gestionnaire des bases

À chaque action de l'utilisateur (souris ou clavier) qui nécessite une modification de la base de connaissances, l'agent éditeur construit le message associé, puis il l'envoie à l'agent gestionnaire des bases de connaissances.

Par exemple, lorsque l'utilisateur déplace la représentation graphique d'une entité, il change ses coordonnées. Cette modification implique donc la mise à jour des attributs « x » et « y » de l'entité associée du réseau sémantique correspondant à la vue de la modélisation courante. On obtient alors un message LSNCW équivalent à celui-ci :

```
(defineValue element="e" modeling="Ma Modélisation" semanticNetwork="Vue #1"
slot="x" value="216")
(defineValue element="e" modeling="Ma Modélisation" semanticNetwork="Vue #1"
slot="y" value="134")
```

Remarque : SNCW propose un module pour afficher l'ensemble des messages envoyés à l'agent gestionnaire des bases de connaissances. Ce module joue le rôle de journal d'évènements.

4.3.5 Intégration dans KCW

Comme décrit précédemment (voir « KCW » page 161), SNCW est un des ateliers de KCW. À ce titre, les bases de connaissances construites doivent être exploitables par les autres ateliers (comme OCW, l'éditeur d'ontologies) mais aussi par des applications tierces. Ainsi, pour favoriser leur réutilisation et leur exploitation, SNCW sauvegarde ses bases dans un langage au format XML. Ainsi, elles sont facilement exploitables par les autres ateliers ou bien par des applications tierces.

SNCW permet aussi de partager ses bases de connaissances avec les différents outils de KCW (les « Os »). Pour chaque outil, il permet d'exporter tout ou partie de la base au format des outils concernés. Ces exportations sont particulièrement utiles pour permettre de cartographier les bases avec d'autres paradigmes.

4.3.6 Synthèse

SNCW est un outil de cartographie sémantique pour construire des modélisations semi-formelles d'un domaine donné. Il exploite d'une part notre méthode de cartographie sémantique mais aussi notre formalisme de description et notre processus. Il permet aux experts de modéliser les connaissances liées à leur domaine avec pour finalité la construction d'un système d'entités structurées capable d'exprimer leurs connaissances. Le résultat est la production d'une « base de connaissances ».

La réalisation de SNCW est basée sur un ensemble d'agents logiciels qui communiquent à l'aide de deux langages dédiés à la manipulation de bases de connaissances (LSNCW) et de réseaux sémantiques (LSN).

SNCW appartient à KCW et à ce titre les bases qu'il permet de produire sont partagées avec les différents ateliers et les outils de KCW. Le format des bases étant ouvert, elles peuvent aussi être partagées avec des applications tierces.

4.4 Exploration sémantique avec EyeTree et RadialTree

L'EyeTree et le RadialTree sont l'implémentation directe de nos paradigmes dédiés à l'exploration sémantique d'espaces informationnels structurés ontologiquement. Ils répondent aux critères de cartographie que nous avons proposés (voir à la page 131) :

- > Utiliser une technique de type « focus + context » pour permettre à l'utilisateur de se concentrer sur certains éléments tout en facilitant l'accès aux autres éléments ;
- > Utiliser une géométrie euclidienne pour ne pas perturber la perception naturelle des manipulations du plan ;
- > Proposer une vue globale de la modélisation (l'ontologie) permettant à l'utilisateur de facilement appréhender l'ensemble de l'espace informationnel ;
- > Pouvoir parcourir la base de connaissances tout en gardant un point fixe de référence.

Ils proposent également une exploration selon deux profils distincts d'utilisateurs : les « novices » et les « experts ». Les « novices » ont une certaine connaissance des concepts du domaine sans connaître exactement leur organisation ; alors que les « experts » ont une bonne maîtrise de l'ontologie de leur domaine. Ces deux profils n'ont pas les mêmes attentes concernant la vision globale de la structure. Pour les novices, il est nécessaire de leur proposer une carte respectant au mieux cette structure pour mieux la partager. En revanche, il est nécessaire de proposer aux experts une carte de la structure pour se situer rapidement dans l'espace informationnel.

4.4.1 Os EyeTree

Le « EyeTree » est notre premier paradigme de visualisation de concepts dédié aux « novices » du domaine. Il utilise une technique de représentation avec déformation de type « Fisheye Polaire ».

L'espace informationnel *représenté* possède une géométrie euclidienne et les nœuds (représentation des concepts) y sont répartis radialement avant de subir une opération de visualisation avec une fonction d'amplification continue appliquée aux coordonnées polaires des nœuds.

a. La fonction d'amplification

Les nœuds sont répartis radialement dans l'espace informationnel *représenté* et une fonction d'amplification est appliquée pour obtenir un effet fisheye. L'utilisateur a alors l'impression que la taille des nœuds et la distance entre chaque nœud sont inversement proportionnelles à leur distance au centre du disque. En d'autres termes, plus les nœuds sont éloignés plus leur taille et leur distance sont réduites.

La fonction d'amplification permet de calculer les coordonnées des nœuds dans la carte à partir de leurs coordonnées dans l'espace informationnel *structuré*. Pour obtenir l'effet fisheye, nous avons besoin d'une fonction qui « ramène l'infini au fini », c'est-à-dire que nous avons eu besoin d'une fonction qui tend vers une valeur fixe lorsque son paramètre tend vers l'infini.

Comme base pour notre fonction d'amplification, nous avons choisi la fonction tangente hyperbolique notée th , définie sur \mathfrak{R} par :

$$th(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{1 - e^{-2x}}{1 + e^{-2x}} \text{ avec } th : \mathbb{R} \rightarrow]-1,1[$$

La fonction tangente hyperbolique est une fonction impaire strictement croissante sur \mathbb{R} :

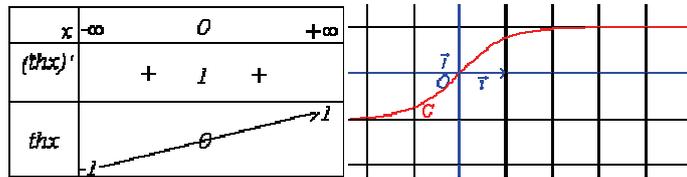


Figure 114 - EyeTree : caractéristiques de la fonction d'amplification.

En prenant la position des nœuds dans l'espace représenté comme paramètre, notre fonction d'amplification permet de les rapprocher du centre de la carte proportionnellement à leur distance de ce centre. Pour obtenir un effet sphérique, nous avons appliqué cette fonction aux coordonnées polaires de chaque nœud de l'espace représenté.

Voici un exemple avec une arborescence répartie radialement :

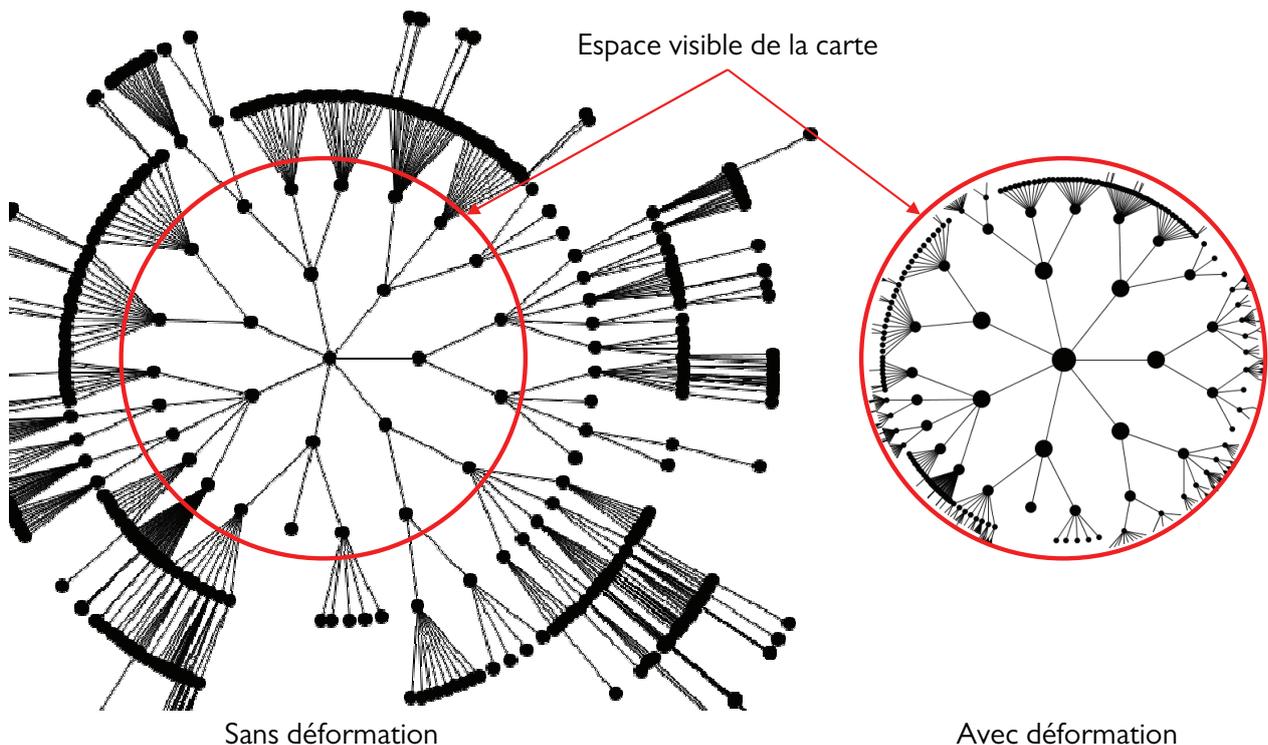


Figure 115 - EyeTree : résultat de l'application de la fonction d'amplification.

Comme nous l'avons précisé dans sa spécification (voir à la page 132), le résultat est similaire aux arbres hyperboliques, mais les interactions de l'utilisateur (par exemple les translations) sont appliquées à un plan euclidien ; comme avec les murs fuyants (voir à la page 83), elles sont « naturellement » prévisibles par l'utilisateur [Leung & Apperley, 1994].

b. Intégration dans KCW

L'EyeTree fait partie des outils de KCW (les « Os » ; voir en page 161). À ce titre, il permet de cartographier les connaissances manipulées par les ateliers et les outils de KCW tel que SNCW.

Les communications entre les autres éléments de KCW et l'EyeTree se font avec un format d'échange et de description d'arborescences de type XML. Ce langage permet de décrire une arborescence d'éléments avec pour chaque élément :

- > Des propriétés graphiques comme la couleur du nœud, la couleur du bord, une image, etc.
- > Des éléments à afficher uniquement dans le panneau latéral (comme des ressources, des sous-arbres, etc.).
- > Un menu contextuel pour associer des actions ou des ressources externes.

En utilisant ce langage, les autres éléments de KCW n'ont alors plus qu'à générer leurs données dans ce format pour les cartographier avec l'EyeTree.

Prenons par exemple le cas de SNCW. Il permet de construire des bases de connaissances basées sur des réseaux sémantiques (voir en page 172). Pour chaque entité de la modélisation, il est possible de générer une arborescence au format EyeTree. Pour cela, SNCW construit un arbre avec l'élément sélectionné comme racine. Puis, il ajoute récursivement comme sous-arbre tous les éléments qui lui sont liés directement ou indirectement (en évitant les cycles).

Voici un exemple d'utilisation de SNCW pour cartographier une arborescence avec l'EyeTree :

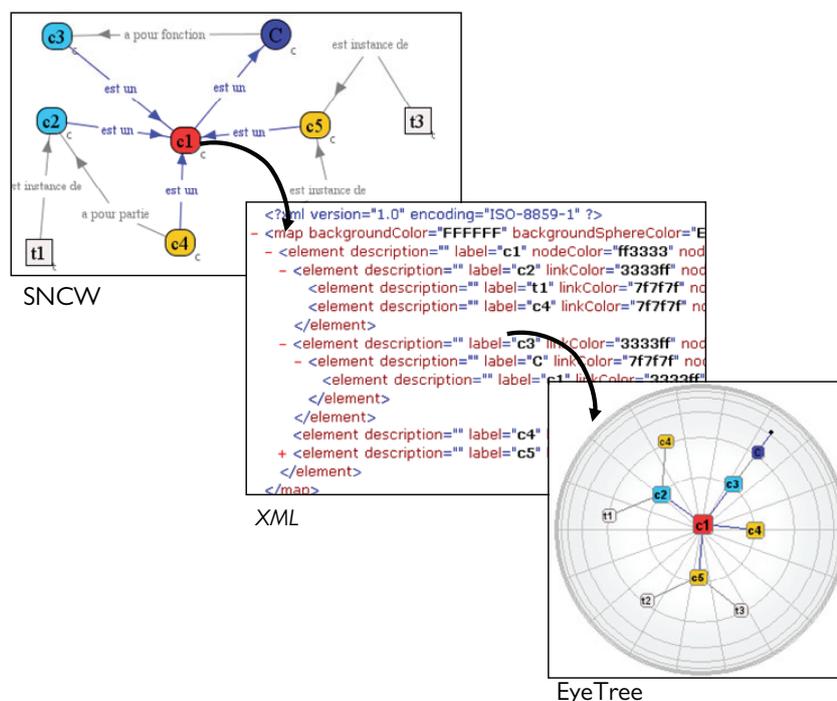


Figure 116 - Exemple d'intégration de l'EyeTree dans KCW.

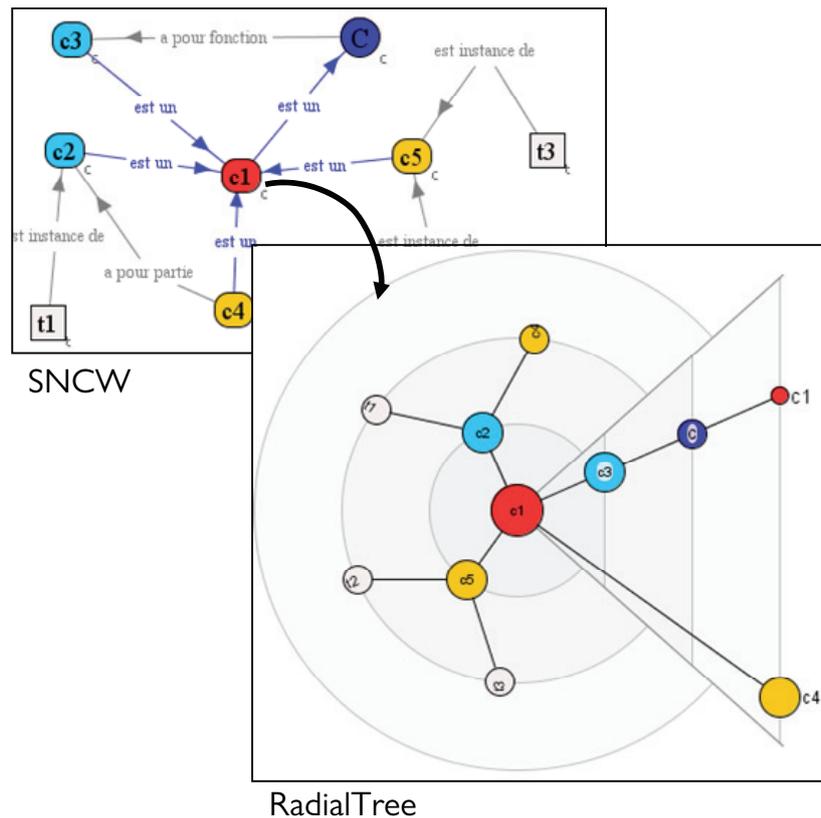


Figure 119 - Exemple d'intégration de RadialTree dans KCW.

4.4.3 Synthèse

L'EyeTree et le RadialTree sont l'implémentation directe de nos paradigmes dédiés à l'exploration sémantique d'espaces informationnels structurés ontologiquement.

Le « EyeTree » est notre premier paradigme de visualisation de concepts dédié aux « novices » du domaine. Il utilise une technique de représentation avec déformation de type « Fisheye Polaire ». Par opposition à l'EyeTree, le « RadialTree » est plus particulièrement dédié aux experts du domaine. Il en reprend les principales caractéristiques comme une cartographie globale de l'ontologie et l'utilisation d'une géométrie euclidienne. Cependant, le RadialTree n'applique pas de déformation globale à l'espace de représentation.

Ces deux nouveaux paradigmes font partie des outils de KCW. À ce titre, ils permettent de cartographier les connaissances manipulées par les ateliers et les outils de KCW (tel que SNCW) en utilisant un format d'échange et de description d'arborescence de type XML.

Bien que ces deux paradigmes soient dédiés à la cartographie d'espace informationnel structurés ontologiquement, l'utilisation de leur langage de description d'arborescence permet de les utiliser pour cartographier tous types d'arborescences.

4.5 Synthèse des réalisations

Cette partie présente l'ensemble des réalisations associées à cette étude. Elles ont été réalisées dans le contexte particulier d'une Équipe de Recherche Technologie (ERT) associée à une société de valorisation leur conférant alors un caractère très industriel.

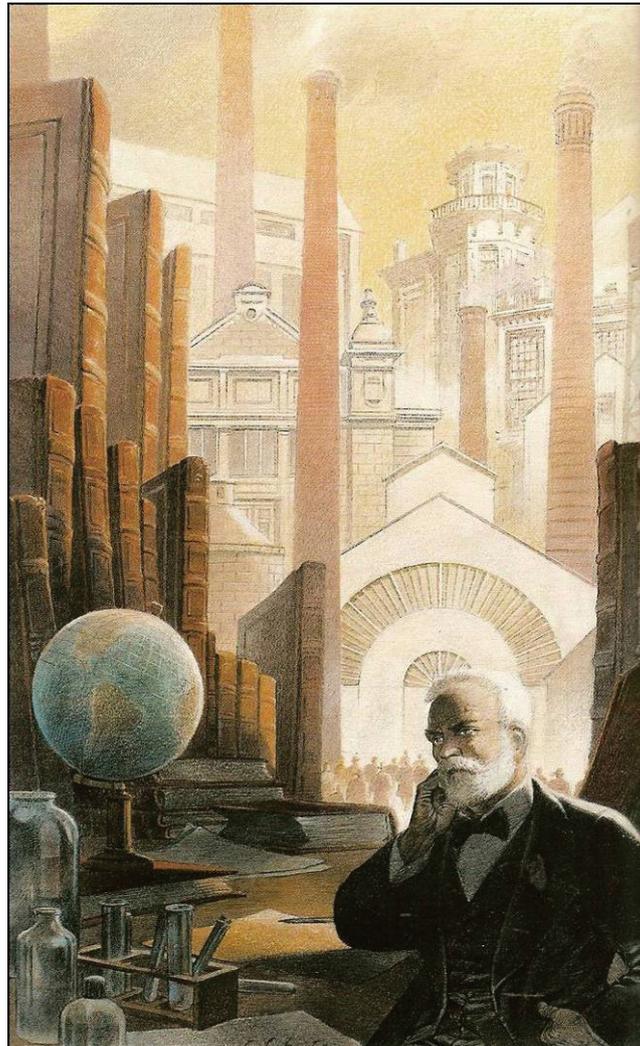
La présentation des réalisations débute par Os Map, un service web de cartographie sémantique mettant directement en œuvre l'ensemble de nos propositions (méthode, processus, modélisation, paradigmes, langage et architecture).

SNCW correspond à notre deuxième réalisation. Il permet d'exploiter notre méthode de cartographie sémantique pour construire une modélisation d'un domaine donné. SNCW met aussi en œuvre notre proposition de modélisations ainsi que notre processus.

Enfin, les deux réalisations présentées sont l'EyeTree et le RadialTree correspondant à l'implémentation de notre spécification de paradigmes pour l'exploration.

La partie suivante présente deux mises en œuvre et validations en gestion des connaissances de nos différentes propositions et réalisations.

Partie 5 : **Mises en œuvre & validations** **en gestion des connaissances**



« Il n'y a rien que l'homme soit capable de vraiment dominer : tout est tout de suite trop grand ou trop petit pour lui, trop mélangé ou composé de couches successives qui dissimulent au regard ce qu'il voudrait observer. Si ! Pourtant, une chose et une seule se domine du regard : c'est une feuille de papier étalée sur une table ou punaisée sur un mur. L'histoire des sciences et des techniques est pour une large part celle des ruses permettant d'amener le monde sur cette surface de papier. Alors, oui, l'esprit le domine et le voit. Rien ne peut se cacher, s'obscurcir, se dissimuler. »

Bruno Latour [Latour, 1985]

Rappel du plan

5.1 Cartographie pour la gestion des compétences.....	199
5.1.1 Problématique.....	199
5.1.2 Notre approche.....	200
5.1.3 Les cartes annuaires.....	202
5.1.4 Les cartes des emplois accessibles.....	203
5.1.5 Les cartes prévisionnelles.....	204
5.1.6 Les cartes stratégiques.....	206
5.1.7 Conclusion : la presse en parle.....	208
5.2 Cartographie pour la navigation dans un fonds documentaire.....	210
5.2.1 Problématique.....	210
5.2.2 Notre approche.....	212
5.2.3 Exploration graphique.....	212
5.2.4 Recherche graphique.....	214
5.2.5 Mise en œuvre : Misti.....	217
5.2.6 Conclusion.....	219
5.3 Synthèse des mises en œuvres.....	220

En accord avec sa fonction d'équipe de recherche technologique, l'équipe Condillac met en œuvre des solutions de gestion de connaissances pour lever différents verrous technologiques. L'appréhension d'un espace informationnel fait partie de ces verrous à lever et la cartographie sémantique – entendue comme notre approche de cartographie de données abstraites ainsi que nos différentes propositions et réalisations – est notre solution.

Cette partie a pour objectif de présenter deux cas concrets de mise en œuvre dans le monde industriel des solutions de gestion de connaissances de l'équipe Condillac et plus particulièrement de notre solution de cartographie sémantique.

Le premier cas correspond à la cartographie sémantique appliquée à la gestion des compétences. Ce cas est à l'initiative de deux organisations : SADEC, un groupe d'expertise comptable et la ville de Cran-Gevrier. Cette étude sur la gestion des compétences a abouti à la réalisation d'Os-Skill, environnement logiciel dédié à la gestion prévisionnelle des emplois et des compétences étendue à la gestion des connaissances (GPECC). Les fonctionnalités d'Os-Skill sont intégrées à un environnement de cartes interactives accessibles par Internet et intranet.

Le second cas correspond à l'utilisation de la cartographie sémantique pour gérer des fonds documentaires. Cette mise en œuvre est à l'initiative du GRETh, une émanation du CEA dédiée à la recherche sur les échangeurs thermiques.

5.1 Cartographie pour la gestion des compétences

Depuis quelques années, l'équipe Condillac travaille sur la problématique de la gestion des connaissances appliquée aux ressources humaines. Nous formulons cette problématique par la question suivante : « comment gérer les compétences d'une organisation ? ».

5.1.1 Problématique

Les années quatre-vingt-dix ont été marquées par l'avènement de la notion de « compétence » dans les différentes fonctions des entreprises [Dejoux, 2000]. C'est durant cette période et dans un contexte de restructurations et de plans sociaux que les démarches de gestion prévisionnelle des emplois et des compétences (GPEC) se sont fortement développées. Aujourd'hui la notion de « compétence » est devenue incontournable au sein de l'entreprise¹.

Au-delà de l'aspect compétence des collaborateurs, les organisations ont également l'impératif de conserver leur capital intellectuel. L'entreprise a besoin de gérer ses compétences mais également ses connaissances – *a fortiori* si l'on considère, même de façon simpliste, qu'une compétence est une connaissance en action –. Les raisons sont nombreuses, citons simplement la gestion de projets et des retours d'expérience ou l'innovation concurrentielle. La frontière entre « la gestion des connaissances » et « la gestion des compétences » est mince et les interactions possibles entre les deux sont importantes. Une entreprise qui disposerait d'un système capable de rassembler les différents supports de la connaissance (fiches projets, fiches de retour d'expérience, description de processus) et de

¹ Différents sigles sont utilisés, du courant « GPEC », à la « GPECC », GPEC étendue à la gestion des Connaissances, en incluant la « GAC » pour « Gestion Anticipée des Compétences ».

faire la corrélation entre les compétences, les acteurs et les documents déjà existants dans l'entreprise, posséderait un réel atout.

Nous avons étudié cette problématique dans un contexte particulier décrit dans la section suivante.

a. Contexte

Cette étude sur la gestion des compétences a été menée en collaboration avec deux organisations : la ville de Cran-Gevrier¹ et le groupe « SADEC »², un groupe d'expertise comptable. Ces deux organisations constituent les deux principaux partenaires de ce projet.

La problématique de la gestion des compétences est identique pour ces deux organisations bien qu'elles ne soient pas issues des mêmes besoins. La ville de Cran-Gevrier emploie environ 400 collaborateurs et son objectif à terme est de placer l'homme au centre de sa politique et plus concrètement, de rémunérer ses collaborateurs en fonction de leurs compétences. Le groupe SADEC emploie environ 200 collaborateurs répartis sur 15 sites situés entre Lille et Bourgoin-Jallieu. L'objectif de ce groupe est de développer les échanges interbureaux et ainsi de valoriser les pôles de compétences existants.

b. Positionnement de la problématique

Leurs objectifs ne sont pas réalisables sans une gestion approfondie des compétences des collaborateurs. Donc, que ce soit à des fins sociales ou économiques, la gestion des connaissances appliquées aux ressources humaines implique d'avoir pour chaque organisation d'une part le détail des compétences de chaque collaborateur, et d'autre part une vision globale de l'ensemble des compétences de l'organisation.

Dans tous les cas, nous sommes confrontés à un problème d'exploration d'espace informationnel constitué d'une part, des informations relatives aux collaborateurs (comme leurs compétences) et d'autre part, des informations liées à l'activité de l'organisation (comme les documents).

5.1.2 Notre approche

Face à cette problématique, nous avons choisi de mettre en œuvre notre démarche de cartographie sémantique ; elle consiste à cartographier un espace informationnel en exploitant la sémantique du domaine auquel il appartient.

Cette démarche implique deux phases : la modélisation du domaine puis la cartographie de l'espace informationnel en exploitant la modélisation.

a. La modélisation

Dans le contexte de gestion des compétences par cartographie sémantique, la modélisation nécessite de pouvoir décrire les concepts du domaine concerné. Nous avons opté pour une démarche ontologique dans la mesure où elle repose sur une modélisation conceptuelle, ici les concepts liés à la GPECC (métiers et compétences) et les expressions métiers de l'entreprise [Roche et al., 2005].

¹ <http://www.ville-crangevrier.com>

² <http://www.sadec-ciagec.com>

Comme décrite précédemment (voir « La modélisation ontologique comme référentiel métier » page 160), la démarche ontologique permet de décrire des modélisations de domaines (les référentiels métiers). Dans le cadre de la gestion des compétences, la modélisation repose sur une définition générique – extensible en fonction des applications – de concepts se structurant sous la forme d'une ontologie.

Pour les métiers, nous utilisons les définitions suivantes :

- > **Poste** (de travail) correspondant à la mise en œuvre de compétences en vue de la réalisation d'une tâche (mission, finalité, objectif). Celle-ci est réalisée par une personne à un moment déterminé et dans un lieu donné de l'organisation ;
- > **Emploi** décrivant la fonction mise en œuvre dans le cadre d'un ou de plusieurs postes de travail est définie en termes de compétences ;
- > **Métier** décrivant un savoir relatif à une même pratique qui se réalise dans les emplois rattachés au métier. Ce savoir est défini en termes de domaine(s) de compétences ;
- > **Filière** (domaine d'activités) regroupant les métiers portant sur un même type d'activités.

Nous procédons de même avec les compétences qui se structurent également sous la forme d'une ontologie selon leur nature et leur degré de généralité (domaines de compétences) ou de spécialisation (sous domaines de compétences, compétences opérationnelles) ; sachant qu'une compétence correspond à la réalisation d'une tâche élémentaire. Le concept de domaine (ou famille de compétences) permet de regrouper les compétences relatives à un même domaine.

b. La cartographie

Pour rappel, notre démarche de cartographie sémantique consiste à exploiter la modélisation pour structurer l'espace informationnel afin de construire des cartes permettant d'aider les utilisateurs à réaliser leurs activités. Les ontologies des métiers et des compétences présentées précédemment constituent une modélisation qui permet de structurer l'espace informationnel des organisations. En effet, toutes les informations de l'organisation liées à la gestion des compétences sont décrites selon cette modélisation et constituent l'espace informationnel *structuré* à cartographier. Il nous reste alors à étudier les besoins des utilisateurs pour spécifier les cartes à construire.

À la suite des différentes phases de réflexions et d'échanges avec les partenaires du projet, nous avons identifié quatre activités nécessitant une aide pour faciliter leur réalisation :

- > Gestion des annuaires : accéder aux collaborateurs par leurs métiers, missions, compétences techniques et réseaux professionnels.
- > Gestion des aires de mobilité : identifier l'ensemble des cheminements professionnels envisageables de chaque collaborateur.
- > Gestion prévisionnelle des emplois et des compétences : suivre l'évolution des métiers, compétences techniques et compétences comportementales dans l'organisation.
- > Analyse stratégique : observer l'organisation selon certaines perspectives comme les formations des collaborateurs.

Chaque carte est dédiée à une activité. Par conséquent, nous avons défini autant de cartes que d'activités : les cartes annuaires, les cartes prévisionnelles, les cartes stratégiques et les cartes des emplois accessibles.

Cette étude sur la gestion des compétences a abouti à la réalisation d'Os-Skill, environnement logiciel dédié à la GPECC. L'ensemble des fonctionnalités d'Os-Skill est intégré à un environnement de cartes interactives accessibles par Internet et intranet. Toutes les cartes mises en œuvre dans Os-Skill sont décrites dans les sections suivantes.

5.1.3 Les cartes annuaires

Un des objectifs de la démarche de gestion des compétences est de valoriser les pôles de compétences existants. Dans ce dessein, il est nécessaire de donner la possibilité d'accéder aux collaborateurs des organisations en fonction de leurs compétences. Ainsi, il devient possible d'identifier le collaborateur le plus qualifié pour trouver une solution à un problème donné.

Pour atteindre cet objectif, nous avons défini un ensemble de cartes « annuaires ». Chacune de ces cartes donne un point d'accès à l'ensemble des collaborateurs selon une perspective particulière : accès par les métiers, missions, compétences techniques ou réseaux professionnels.

Les cartes annuaires exploitent l'EyeTree, notre paradigme présenté précédemment (voir « Os EyeTree » page 189). Le référentiel (l'ontologie) est cartographié avec l'EyeTree et lorsqu'une compétence est sélectionnée, un volet latéral affiche l'ensemble des collaborateurs qui possèdent cette compétence.

La saisie d'écran ci-après présente un exemple de carte annuaire sur les compétences techniques¹:

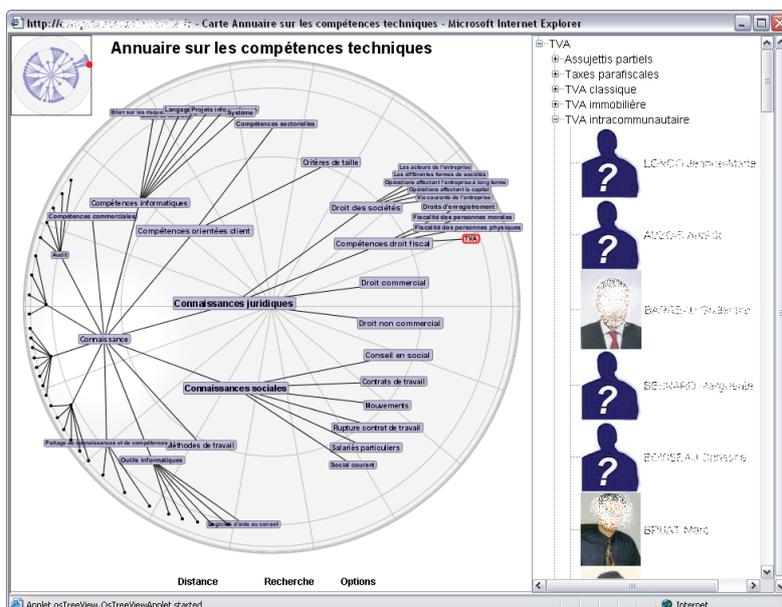


Figure 120 - Exemple de carte annuaire.

¹ Pour des raisons de confidentialité certaines informations graphiques sont altérées.

Les cartes annuaires permettent aux collaborateurs d'identifier la personne correspondant à leurs besoins selon ses compétences. Cette cartographie guidée par les référentiels métiers (la modélisation) permet aux collaborateurs d'appréhender leur organisation comme un système de compétences détenues par des personnes et non comme un ensemble de personnes où chaque personne est attachée à un emploi.

5.1.4 Les cartes des emplois accessibles

La gestion des aires de mobilité permet d'identifier l'ensemble des cheminements professionnels envisageables de chaque collaborateur. L'objectif est alors de permettre aux collaborateurs de formuler des souhaits d'évolution de carrière en tenant compte de leurs compétences indépendamment des postes existants ou à venir. *In fine*, les collaborateurs peuvent souhaiter une poursuite de carrière vers ce que nous nommons les « emplois accessibles ».

Chaque collaborateur possède un ensemble de compétences avec un niveau évalué (parmi base « - », qualifié « -+ », maîtrise « + » et expert « ++ ») et chaque emploi nécessite de posséder un ensemble de compétences à un niveau donné. La recherche des emplois accessibles permet d'identifier tous les emplois accessibles pour un collaborateur donné, moyennant éventuellement l'acquisition ou le perfectionnement de certaines compétences.

Concrètement pour un collaborateur et un emploi donné, le collaborateur peut se situer selon trois cas :

Cas 1 : le collaborateur possède au minimum toutes les compétences requises pour le poste avec au moins le niveau requis.

Cas 2 : le collaborateur possède toutes les compétences requises, mais il ne possède pas toujours le niveau requis. Dans ce cas, il lui est nécessaire de suivre des formations pour perfectionner les compétences incriminées.

Cas 3 : le collaborateur ne possède que certaines compétences requises (éventuellement avec des compétences à perfectionner). Dans ce cas, il doit acquérir celles qui lui manquent avec le niveau requis.

La carte des emplois accessibles exploite cette classification en cas pour représenter les emplois accessibles. La carte est composée d'un demi-disque, lui-même décomposé en trois demi-disques superposés à raison d'un disque par cas (voir la figure suivante). Le collaborateur concerné est représenté par sa photo et son nom placés au centre du demi-disque. Chaque emploi est représenté par un petit disque de couleur. Tous les emplois correspondant au même cas, sont répartis sur l'extrémité du demi-disque correspondant.

Voici un exemple de carte réalisé pour le collaborateur « Dupond Thierry » ; le cercle rouge indique l'emploi sélectionné et le détail de cet emploi est affiché dans le volet de droite :

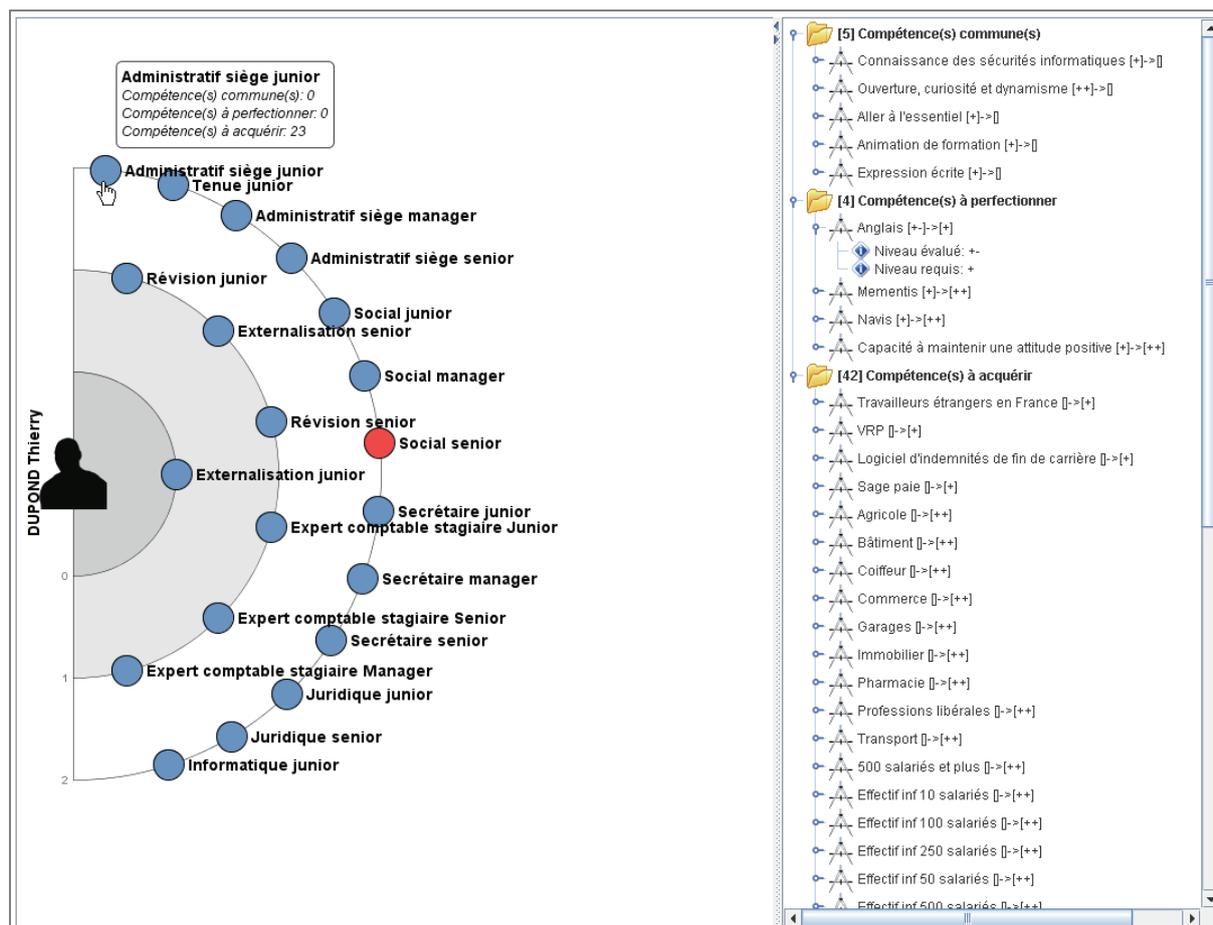


Figure 121 - Exemple de carte des emplois accessibles.

La représentation choisie suggère que plus un emploi est proche d'un collaborateur (proche de sa photo) plus il lui est facilement accessible et inversement. La distance d'une représentation d'un emploi et de la photo est donc proportionnelle à l'effort requis (calculé et supposé) pour obtenir cet emploi.

Par exemple, les emplois du cas 3 (ceux qui nécessitent des compétences à acquérir) sont placés sur le demi-cercle le plus éloigné de la photo et les emplois du cas I (ceux qui ne nécessitent aucun perfectionnement et aucune acquisition) sont placés sur le demi-cercle le plus proche de la photo.

Lorsqu'un emploi est sélectionné, un volet latéral affiche le détail complet sur l'accessibilité de cet emploi avec : la liste des compétences communes, la liste des compétences à perfectionner et la liste des compétences à acquérir.

Le survol d'un emploi avec le curseur de la souris (comme dans la figure précédente) permet d'afficher une bulle d'information synthétique sur cet emploi. La carte permet aussi aux collaborateurs de soumettre directement leurs souhaits de mobilité au service des ressources humaines ou bien encore, d'afficher le détail des formations possibles pour toute acquisition ou perfectionnement de compétences.

5.1.5 Les cartes prévisionnelles

Depuis quelques années maintenant, les organisations et plus particulièrement les entreprises prennent conscience de la valeur de leurs savoir-faire. La mise en œuvre de

méthodes de gestion de connaissances est la conséquence directe de cette prise de conscience.

La problématique majeure est alors de savoir comment préserver ce capital. Ces savoir-faire dont il est question, sont détenus par des collaborateurs qui possèdent eux-mêmes les compétences participant à ces savoir-faire. Chaque collaborateur possède un contrat de travail qui le lie à l'organisation et par conséquent, il est possible d'estimer pour combien de temps il peut apporter ses compétences. Grâce aux référentiels métiers et l'évaluation des compétences des collaborateurs, il devient aussi possible de savoir le nombre de personnes qui détiennent une compétence donnée et pour combien de temps. De plus, la prise en compte des missions permet de connaître l'utilisation des compétences et par voie de conséquence, de planifier leurs utilisations.

Les cartes prévisionnelles ont été conçues pour répondre à la problématique de la préservation du savoir-faire en permettant de se projeter dans l'avenir est d'observer l'organisation selon la perspective de ses compétences. Elles permettent de suivre l'évolution des métiers, des compétences techniques et des compétences comportementales dans l'organisation.

Ainsi, les cartes permettent de visualiser l'ensemble des compétences de l'organisation et ainsi de mettre en valeur celles qui sont amenées à disparaître ou à manquer (nommées les « compétences critiques »), ou au contraire celles qui vont être en surnombre.

Les cartes prévisionnelles exploitent une variante de notre paradigme RadialTree décrit précédemment (voir « Os RadialTree » page 192) pour cartographier le référentiel concerné. Les domaines de compétences sont représentés par des cercles. La taille des cercles est proportionnelle au nombre de compétences du domaine. Ainsi, il est possible d'identifier rapidement l'importance des domaines de compétences. La couleur est fonction du nombre de compétences critiques dans la projection : en rouge si le domaine possède des compétences critiques et en gris sinon.

Si un domaine est sélectionné, un volet latéral affiche les compétences du domaine ainsi que leur état (nombre de personnes à posséder la compétence et éventuellement pourquoi elle est critique).

Par exemple, la figure ci-dessous représente le référentiel des compétences techniques d'une organisation projetée dans 5 ans :

Par exemple, la saisie d'écran ci-après¹ correspond à une pyramide des âges des collaborateurs pour différents bureaux. Chaque barre correspond à une tranche d'âge ; la taille des barres représente l'effectif des collaborateurs ; les barres sont regroupées par bureau et par ordre de tranche d'âge.

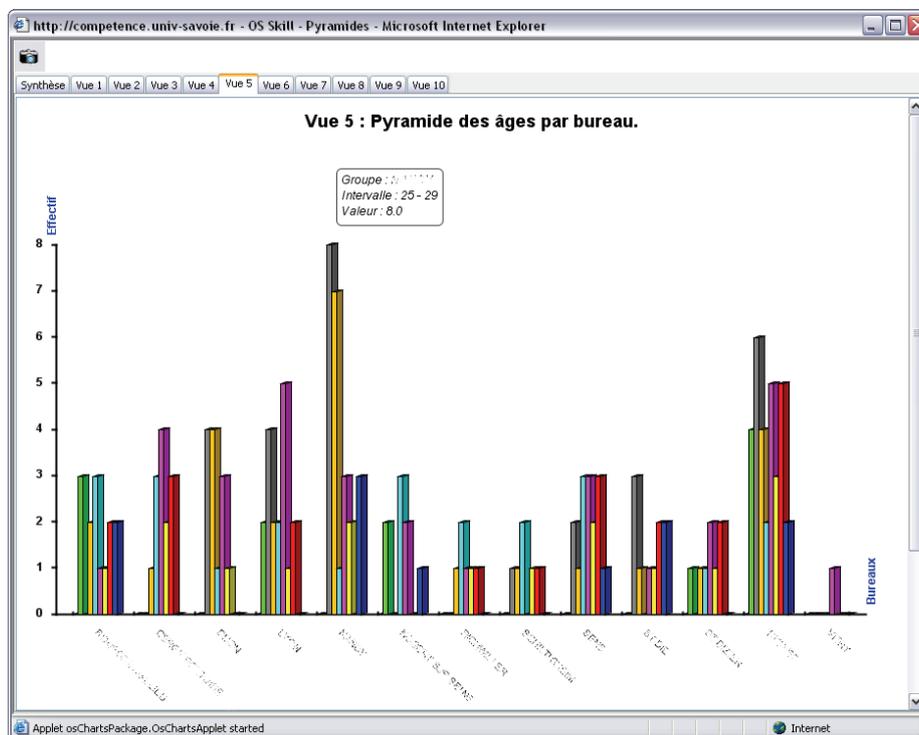


Figure 123 - Carte stratégique : pyramide des âges des collaborateurs par bureau.

Les barres correspondant à chaque tranche d'âge dans les différents bureaux possèdent la même couleur ce qui facilite leur comparaison. Le survol avec la souris d'une barre permet d'en afficher le détail dans une bulle comme dans l'exemple précédent.

Toutes les cartes stratégiques d'une même organisation sont regroupées en une seule carte composée d'onglets à raison d'un onglet par carte et d'un onglet supplémentaire regroupant en miniatures toutes les cartes.

Voici un exemple de l'onglet des miniatures :

¹ Pour des raisons de confidentialité certaines informations graphiques sont altérées.

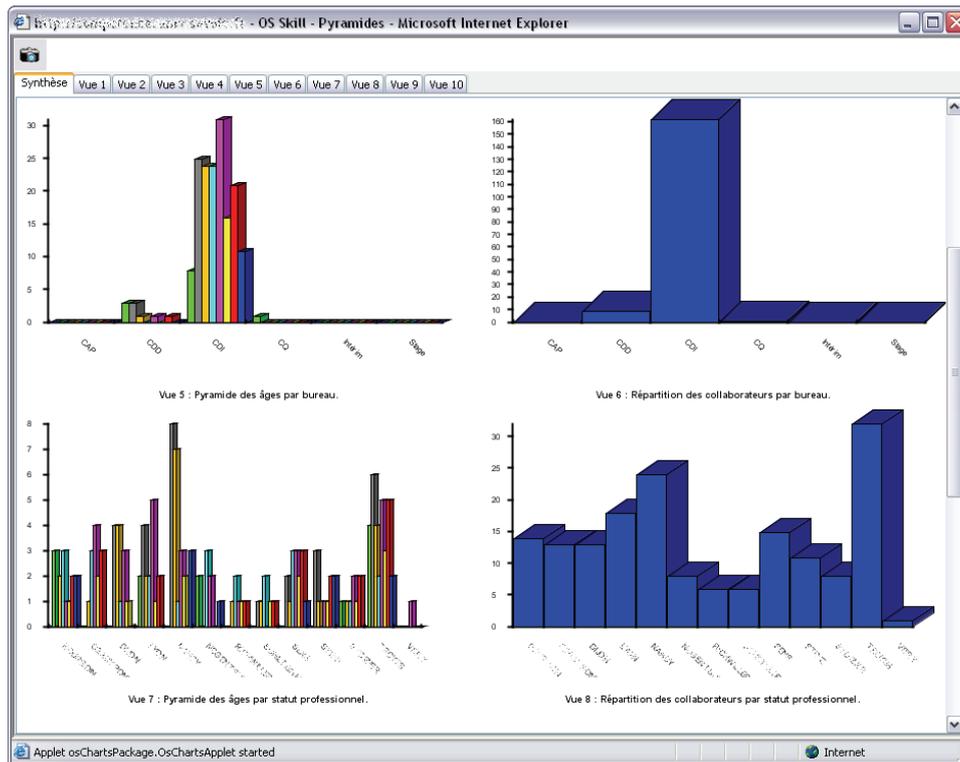


Figure 124 - Cartes stratégiques : vision globale.

Après le déploiement de cartes stratégiques, les utilisateurs ont exprimé le besoin de pouvoir imprimer les cartes pour servir de document de références lors de réunions selon le principe des tableaux de bord. Nous avons donc rajouté cette fonctionnalité.

5.1.7 Conclusion : la presse en parle

Os-Skill et plus particulièrement la cartographie sémantique des compétences, a fait l'objet d'un article dans la revue « Entreprise & Carrières », l'hebdomadaire de référence sur les ressources humaines, du 7 mars 2006.

Entreprise & Carrières L'HEBDOMADAIRE DES RESSOURCES HUMAINES

Sadeq Une démarche orientée "business"

Sadeq, un groupe d'expertise comptable de 200 salariés, mise sur une technologie cartographique de compétences pour mieux répondre aux besoins de ses clients.

SADEQ

- > **Activité :** expertise comptable.
- > **Effectifs :** 200 salariés.
- > **Chiffre d'affaires :** 11 millions d'euros en 2005.
- > **Solution informatique :** Ontologs.

IE n novembre 2005, les 200 collaborateurs de Sadeq, groupe d'expertise comptable composé de quinze bureaux en France, ont passé un entretien d'évaluation annuel nouvelle formule. La grille d'évaluation des compétences était, pour la première fois, réellement adaptée aux spécificités de chacun des neuf métiers identifiés dans l'entreprise. En outre, une distinction entre junior, senior et manager était introduite.

Grille informatisée

Autre nouveauté, la grille était informatisée. « Nous ne disposions pas, jusqu'alors, de référentiel des compétences. La grille qui préexistait était très générique », reconnaît Thierry Galland, directeur technique de Sadeq et expert-comptable associé. Il pilote, actuellement, un programme de gestion des compétences, dont la première phase consistait justement à enregistrer des informations qualifiées issues des entretiens d'évaluation.

01 - du 7 au 13 mars 2006 **Entreprise & Carrières** 27

Figure 125 - Entreprise & Carrières : extraits.

Dans ce dossier, Thierry Galland, un des responsables du groupe SADEC, témoigne et il explique que la solution cartographique d'Os-Skill leur permet de mieux répondre aux besoins de leurs clients.

5.2 Cartographie pour la navigation dans un fonds documentaire

Cette section présente notre mise en œuvre de la cartographie sémantique pour la navigation dans un fonds documentaire composé de documents numériques scientifiques et techniques.

Notre objectif était de fournir un système permettant à des collaborateurs d'un même domaine d'identifier et d'accéder à l'information nécessaire à leur activité ; les informations sont contenues dans leurs documents.

Cette étude a été initiée avec le « Groupement pour la Recherche sur les Échangeurs Thermiques » (GRETh) qui a mis en place il y a deux ans un site Internet¹ pour la diffusion des connaissances et des informations scientifiques et techniques au service de ses adhérents (principalement des industriels). Une base de données a alors été constituée. Elle contient plus de cinq cents documents dont des rapports de thèses, des rapports scientifiques et techniques (issus de recherches effectuées dans le cadre contractuel de la recherche collective du laboratoire GRETh), de la bibliographie scientifique (issue de publications d'articles de différentes sources), ainsi que de nombreux autres documents. Tous ces documents se rapportent aux métiers du GRETh qui sont basés sur la mécanique des fluides et la thermique des échangeurs. Une phase préliminaire de veille stratégique, menée par le GRETh, a conduit en partie à la constitution de cette base de données. Les utilisateurs qui viennent rechercher de l'information correspondent à une cinquantaine de chercheurs et au moins une centaine d'entreprises industrielles (les collaborateurs des entreprises ayant eux-mêmes accès au système).

Notre approche face à cette problématique a débuté par l'étude des différents paradigmes existants permettant d'accéder à un fonds documentaire scientifique et technique. Notre étude s'est alors rapidement orientée vers les techniques de visualisations interactives pour l'accès et la recherche d'informations annotées sémantiquement. Pour parvenir à notre objectif de cartographie sémantique du fonds documentaire, il nous a fallu au préalable mettre en place un moteur d'indexation automatique s'appuyant sur la sémantique du domaine. Puis, nous avons proposé des cartographies exploitant cette sémantique pour permettre aux usagers d'accéder aux documents. Le résultat de cette approche est la création et la mise en place d'un outil pérenne, car il favorise la diffusion et l'échange de connaissances.

5.2.1 Problématique

a. La documentation scientifique et technique

Un fonds documentaire scientifique et technique est constitué d'un ensemble de documents qui portent sur un même domaine. Au domaine est associée une terminologie dont la signification des termes est définie en relation avec les concepts métier – c'est-à-dire l'ontologie du domaine – et à laquelle est attaché un ensemble de « mots d'usage » utilisé pour la rédaction du corpus.

La prise en compte des mots d'usage permet de gérer la diversité langagière – communauté de pratique et langues naturelles – tout en se référant à la même ontologie du domaine. Par

¹ <http://www.greth.fr>

exemple, les mots d'usage « perte de charge » et « pressure drop » se réfèrent au même concept.

b. Les besoins

Notre objectif est de définir et mettre en œuvre les moyens nécessaires pour accéder à un fonds documentaire, non pas en fonction de mots-clés présents dans les documents mais, selon la modélisation du domaine définie par les experts en termes de concepts métier. L'accès à l'information doit également gérer le multilinguisme car des documents différents dans les termes qu'ils emploient (parce que produits par des communautés de pratiques différentes) peuvent néanmoins se référer aux mêmes notions ; c'est en particulier le cas lorsque les documents sont rédigés dans des langues différentes.

Nous avons pu identifier cinq besoins à satisfaire pour permettre de répondre à cet objectif :

Identifier l'information à l'aide des concepts métier : l'identification de la « bonne information » requiert une recherche basée sur les concepts associés au domaine. Ces concepts sont dénotés par les mots d'usage, termes et expressions, contenus dans les documents. Cette approche de recherche est donc extralinguistique ce qui la rend très différente d'une approche par mots-clés ou par ensemble de mots (sémantique distributionnelle). De plus, une approche par concepts permet d'avoir une indexation elle aussi extralinguistique des documents sur l'ontologie du domaine, garantissant ainsi de retrouver l'information dans les documents indépendamment de leur langue.

Avoir une vision globale du fonds documentaire : permettre à l'utilisateur de s'approprier l'information issue du fonds documentaire nécessite de lui fournir les moyens pour l'appréhender dans sa globalité indépendamment de la diversité langagière des communautés de pratiques. C'est pourquoi, il faut fournir une vision globale de l'ensemble du fonds documentaire en relation directe avec les concepts qu'il manipule.

Naviguer selon la modélisation du domaine : comme décrite précédemment, l'information doit être identifiée à l'aide des concepts métier du domaine. De même, l'accès aux documents doit se faire selon la structuration de ses concepts. Ainsi, la navigation s'effectue selon la modélisation du domaine ce qui permet à l'utilisateur de comprendre, assimiler et exploiter cette modélisation.

Assister l'utilisateur dans l'expression de ses besoins : une des premières difficultés, lors d'une recherche d'informations, est d'exprimer sa requête. Même si l'utilisateur maîtrise bien le domaine, il lui est souvent difficile de choisir les bons éléments pour exprimer ses besoins. Ce problème est particulièrement flagrant lors d'une recherche par mots-clés avec un moteur de recherche. Il faut donc assister l'utilisateur pour construire sa requête en exploitant les connaissances du domaine et leur organisation.

Présenter les résultats : la compréhension des résultats de sa recherche est une autre difficulté pour l'utilisateur. C'est pourquoi, il est nécessaire de fournir une présentation des résultats qui soit adaptée à l'utilisateur et à son activité.

c. Positionnement de la problématique

L'ensemble de ces besoins est lié à la nécessité pour l'utilisateur d'évoluer dans un espace informationnel souvent important. Cette problématique est donc une problématique de navigation et de recherches d'informations dans un espace informationnel constitué du fonds documentaire.

5.2.2 Notre approche

Face à cette problématique, nous avons choisi de mettre en œuvre notre démarche de cartographie sémantique ; elle consiste à cartographier le fonds documentaire en exploitant la sémantique du domaine auquel il appartient.

Cette démarche se déroule en deux étapes : la structuration du fonds documentaire puis sa cartographie.

a. La structuration du fonds documentaire

Pour rappel, notre démarche de cartographie sémantique consiste à exploiter la modélisation pour structurer l'espace informationnel afin de construire des cartes permettant d'aider les utilisateurs à réaliser leurs activités. Dans ce dessein, nous avons appliqué la méthode Ousia qui permet d'obtenir une modélisation ontologique du domaine (Voir « La démarche Ousia » page 160).

Grâce à cette modélisation, la structuration du fonds documentaire correspond à une classification des documents sur l'ontologie du domaine. Elle se fonde sur l'annotation sémantique des documents [Kiryakov *et al.*, 2003], c'est-à-dire sur l'association pour chaque document des concepts référencés par son contenu. Pour cela il est nécessaire de définir au préalable la terminologie de l'entreprise, c'est-à-dire d'associer aux mots d'usage de ses différents vocabulaires (correspondants aux différentes communautés de pratique et langues utilisées) leurs significations en référence aux concepts métier de l'ontologie. L'analyse linguistique du contenu des documents, sur la base des mots d'usage apparaissant dans les textes, permet de les indexer sur les concepts appropriés.

Il est à souligner que la classification obtenue est extralinguistique. Les documents, quelles que soient leurs langues d'écriture et quelles que soient les communautés de pratique qui les produisent, sont indexés sur la même ontologie. Ainsi une requête exprimée dans une langue donnée retournera tous les documents, quelles que soient leurs langues, correspondant aux concepts associés à la requête.

b. La cartographie sémantique

Une fois l'espace du fond documentaire structuré, nous devons le cartographier. La difficulté est alors de déterminer les paradigmes à mettre en œuvre. Face à cette problématique nous avons étudié les différents paradigmes qui permettent de cartographier un fonds documentaire scientifique et technique.

L'étude des besoins a permis de mettre en évidence la nécessité de proposer deux approches cartographiques complémentaires pour naviguer dans le fonds documentaire : par exploration et par recherche. Les sections suivantes décrivent l'étude de chacune de ces approches.

5.2.3 Exploration graphique

Depuis des années, les travaux portant sur l'exploration graphique d'un fonds documentaire ont permis de produire de nombreuses solutions et techniques très variées. Pour déterminer si une de ces solutions pouvait répondre (en totalité ou en partie) à notre problématique, nous avons étudié leurs propriétés.

Face au nombre important de solutions existantes, nous les avons tout d'abord réparties en deux familles : les solutions de visualisation intra-document et les solutions de visualisation

d'une collection de documents. Cette répartition est déjà proposée dans plusieurs travaux [Card *et al.*, 1999a; Jacquemin *et al.*, 2005]. Par exemple, pour illustrer le choix de cette répartition, prenons deux systèmes très représentatifs de ces deux familles : WebBook et DocCube. WebBook [Card *et al.*, 2004; Card *et al.*, 1996] est un système qui permet de représenter des documents avec une métaphore de livre en trois dimensions. DocCube quant à lui [Mothe *et al.*, 2003], représente un fonds documentaire avec des nuages de sphères où chaque sphère correspond à un ensemble de documents.

Dans notre cas, celui d'un fonds documentaire, ce qui nous concerne est la représentation d'une collection de documents. Nous avons donc poursuivi nos travaux par l'étude des solutions de la deuxième famille.

Nous avons ensuite scindé en deux l'ensemble des solutions de visualisations de collections de documents en prenant comme critère discriminant la représentation explicite ou non de la modélisation qui a permis de structurer le fonds documentaire. Il est important de souligner que le fonds documentaire est nécessairement structuré. La modélisation joue alors le rôle de structure pour le fonds documentaire, une sorte de squelette. Cette distinction est importante par son impact sur l'utilisateur. Si elle est explicitement représentée, alors elle est plus facilement appréhendée et mieux assimilée par les utilisateurs. Dans le cas d'une représentation explicite de la structure, l'accès aux documents se fait par l'intermédiaire de leurs classes d'appartenance (indépendamment de la façon dont ont été construites les classes, sémantique distributionnelle ou propre au domaine). En revanche, dans le cas où la structure ne serait pas représentée explicitement, l'accès aux documents est direct.

Pour illustrer cette différence, prenons comme exemple le système Lighthouse [Leuski & Allan, 2000] pour lequel la représentation de la structure n'est pas explicite. Les documents sont représentés par des structures visuelles dans un espace à trois dimensions.

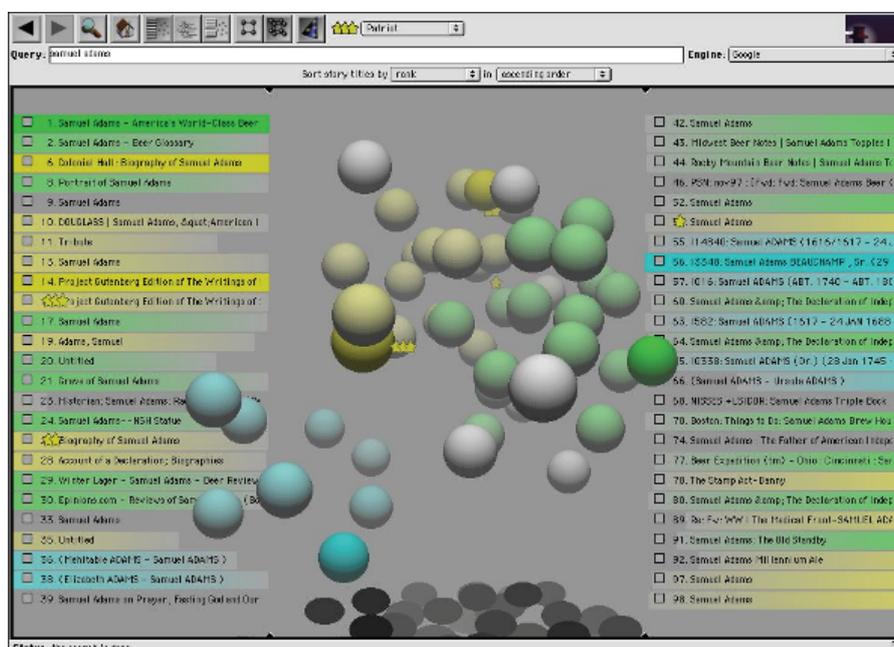


Figure 126 - Le système Lighthouse.

La structuration du fonds documentaire concerné repose sur une fonction de distance (similarité entre les documents) qui donne une répartition spatiale des documents. L'utilisateur

appréhende l'ensemble des documents à travers leur répartition dans la carte. Le système DocCube décrit précédemment appartient aussi à cette catégorie.

Un des besoins lié à notre problématique est d'avoir une exploration du fonds documentaire guidée par la sémantique du domaine, sa modélisation. Par conséquent, il est important d'aider les utilisateurs à la comprendre et à l'assimiler.

Selon notre démarche de cartographie sémantique, nous écartons les solutions qui ne représentent pas explicitement la sémantique du domaine (la modélisation qui permet de le structurer) comme Data Mountain [Robertson *et al.*, 1998], ThemeScapes [Wise *et al.*, 1995], Galaxy Of News [Rennison, 1994], Bib3D [Cubaud *et al.*, 1998; Cubaud & Topol, 2001] et Vibe System [Olsen *et al.*, 1993]. De même, nous avons écarté toutes les solutions exploitant des techniques par pavage (voir « Représentations par pavage » page 72).

À ce stade de l'étude, il reste encore toutes les solutions qui permettent de cartographier une collection de documents en représentant explicitement leur structure.

D'après notre approche, le choix d'une technique de visualisation dépend de la modélisation du domaine. Dans notre cas, la modélisation est une arborescence de concepts. Par conséquent, pour la problématique de l'exploration dans le fonds documentaire, nous avons utilisé l'EyeTree (voir en page 189).

5.2.4 Recherche graphique

Notre objectif est également de proposer une solution de recherche pour permettre aux utilisateurs d'accéder à l'information pertinente pour leur activité. Pour cela, il faut leur fournir les outils pour construire une requête, puis la soumettre au système et enfin, lui présenter les résultats.

Les utilisateurs sont habitués à effectuer des recherches d'informations (l'exemple le plus courant est la recherche sur Internet). Généralement ces recherches sont proposées via une interface possédant un champ de saisie, pour saisir dans un premier temps une requête en langage naturel et une zone pour afficher dans un deuxième temps les résultats.

La requête doit exprimer les besoins de l'utilisateur et tout comme la structuration du fonds documentaire, elle doit s'appuyer sur les connaissances du domaine. Par exemple, l'utilisateur doit pouvoir formuler une requête pour retrouver tous les documents traitant d'un thème donné.

Une des principales difficultés pour les utilisateurs est de construire leurs requêtes et de déterminer les éléments qui doivent la constituer [Mothe *et al.*, 2003]. Un utilisateur même familier avec le domaine concerné saura reconnaître un concept à partir de termes qui le dénotent alors qu'il ne saura pas toujours donner les termes qui dénotent un concept. Nous sommes généralement dans le deuxième cas où l'utilisateur sait quels sujets (thèmes ou concepts) il recherche sans nécessairement connaître tous les mots d'usage utilisés dans le corpus. Dès lors, il apparaît nécessaire d'aider l'utilisateur à construire sa requête en proposant de sélectionner des concepts métier parmi ceux déjà identifiés dans le domaine.

La construction de la requête ainsi que l'exploitation des résultats associés ne doit pas surcharger cognitivement l'utilisateur pour le laisser se concentrer sur sa tâche principale : exploiter la bonne information. Mais, face à un grand nombre de concepts métier, l'utilisateur ne peut plus sélectionner directement ceux qui vont constituer sa requête sans entraîner une surcharge cognitive. Pour permettre une économie cognitive, l'approche courante est d'utiliser des techniques de visualisation d'informations représentant graphiquement les classes du fonds

documentaire. On retrouve cette approche dans le système OntoExplo [Hernandez, 2005; Hernandez & Aussenac-Gilles, 2004] : un treillis de concepts est représenté et lorsqu'un utilisateur sélectionne un concept, les documents associés sont affichés.

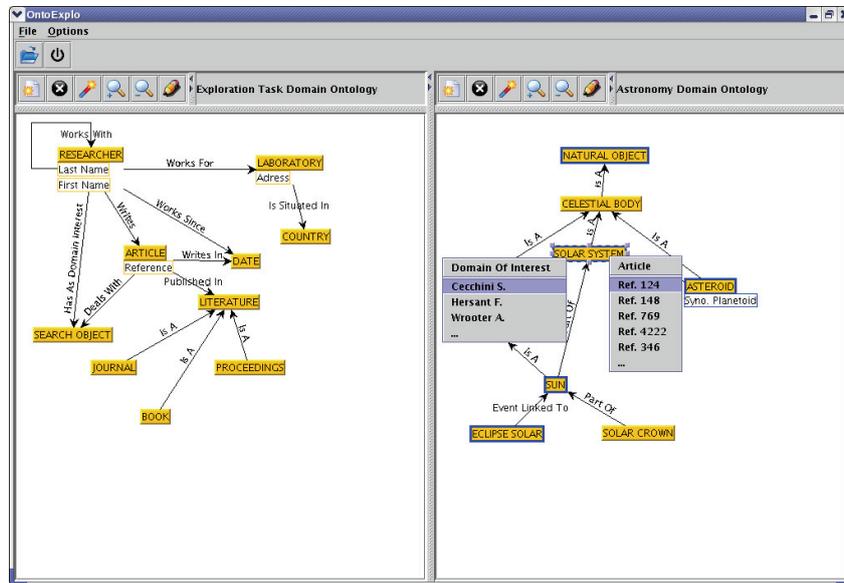


Figure 127 - Recherche de documents avec OntoExplo.

Face à cette problématique de recherche graphique, nous avons défini un nouveau paradigme : OntoRequest.

OntoRequest est un paradigme dédié à la recherche ontologique graphique. Pour cela, il cartographie la requête par un disque au centre de la carte. En périphérie de ce disque sont représentés, par des disques plus petits, les concepts sélectionnés (initialement les racines des différents arbres constituant l'ontologie).

La figure suivante présente une requête constituée de cinq concepts : « Condition de fonctionnement », « Fluide », « Objectif », « Mode de Fabrication » et « Transfert »).

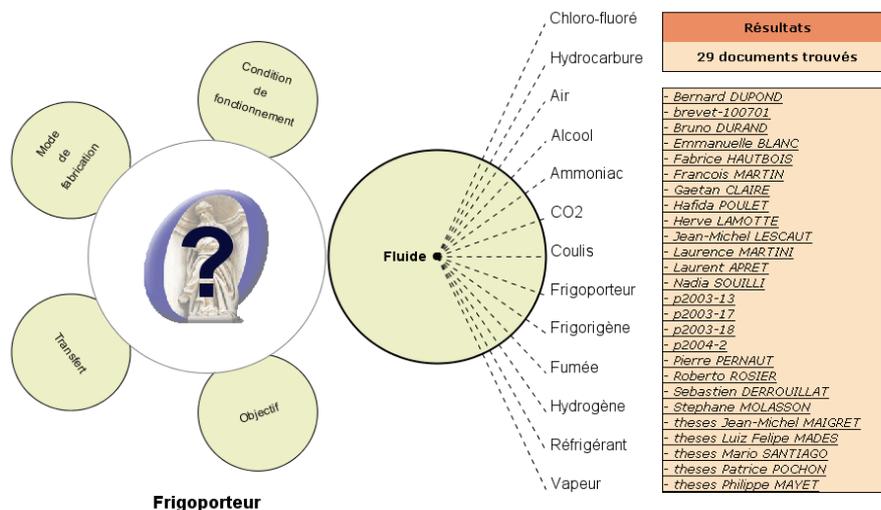


Figure 128 - OntoRequest : construction de requêtes ontologiques

Si l'utilisateur souhaite modifier sa requête afin de l'étendre ou de la restreindre, il lui suffit de spécialiser (ou généraliser) les concepts en parcourant la relation de généralisation/spécialisation. Pour chaque concept à spécialiser (ou à généraliser), l'utilisateur sélectionne sa représentation qui se place dynamiquement le long de la périphérie du disque central, à droite de la figure pour déployer ses sous-concepts. Il suffit ensuite à l'utilisateur soit, de sélectionner un concept-fils pour affiner sa recherche soit, de remplacer (par double clic) le concept lui-même par son concept-père pour étendre la recherche. Le résultat de la requête, représenté par le disque central, est l'intersection des ensembles de documents rattachés aux concepts situés en périphérie sachant qu'à chaque concept est associée l'union des documents rattachés à tous ses concepts-fils.

Voici ci-dessous une saisie d'écran d'une requête constituée de cinq concepts (« Condition de fonctionnement », « Fluide », « Objectif », « Mode de Fabrication » et « Transfert ») :

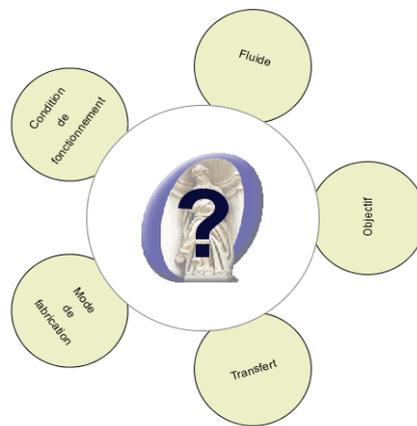


Figure 129 - OntoRequest : exemple de requête.

La saisie d'écran ci-dessous représente le résultat de la sélection du concept « fluide » :

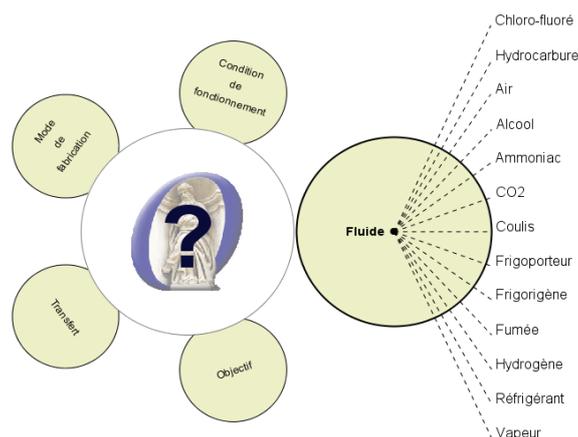


Figure 130 - OntoRequest : sélection d'un concept.

Le résultat de la requête, représenté par le disque central, est l'intersection des ensembles de documents rattachés aux concepts situés en périphérie (à chaque concept est associée l'union des documents rattachés à tous les concepts qu'il subsume).

5.2.5 Mise en œuvre : Misti

Le « Groupement pour la Recherche sur les Échangeurs Thermiques » (GRETh) – à l'origine de ces travaux – a mis en place un site Internet pour la diffusion des connaissances et des informations scientifiques et techniques au service de leurs adhérents qui sont principalement des industriels. Ces informations (articles, thèses, rapports techniques et scientifiques, etc.) sont regroupées au sein d'une base de données. Tous ces documents se rapportent aux métiers du GRETh, basés sur la mécanique des fluides et la thermique des échangeurs.

Dans ce contexte-là, notre approche a été mise en œuvre dans le système MISTI (Moteur Internet Sémantique pour la Thermique Industrielle) : portail web de recherche et de navigation dans un fonds documentaire scientifique et technique.

En tenant compte des résultats présentés précédemment, nous avons défini un système offrant à l'utilisateur trois approches pour retrouver l'information dont il a besoin pour son activité :

- > Une approche classique : l'interface principale.
- > Une approche par exploration : l'EyeTree.
- > Une approche par construction de requête : l'OntoRequest.

Dans le système Misti, les trois approches sont accessibles par des onglets : un onglet par approche de recherche. Elles sont décrites dans les parties suivantes.

a. Recherche classique

Les utilisateurs sont habitués à effectuer des recherches d'informations en soumettant à un système une requête en langage naturel. Nous estimons que cette méthode de recherche « classique » est nécessaire pour ne pas perturber les utilisateurs habitués à cette méthode. C'est pourquoi, MISTI propose une méthode de recherche classique mais, tout comme l'EyeTree et l'OntoRequest, en exploitant l'ontologie du domaine.

La recherche est effectuée en deux temps : la construction de la requête et l'affichage des résultats.

La construction de la requête. L'utilisateur exprime sa requête en langage naturel dans un champ de saisie. Elle est constituée de termes qui dénotent pour l'utilisateur des concepts métier. La requête est ensuite soumise au système qui analyse les termes, identifie des concepts métier et retourne les documents associés.

Par exemple, l'utilisateur va demander au système de trouver « les documents qui traitent des conditions de fonctionnement ». Le système identifie alors le concept « Condition de fonctionnement » et retourne les documents associés.

Pour assister l'utilisateur, le système propose une vue arborescente classique de la hiérarchie de concepts métier du domaine. Lorsque l'utilisateur sélectionne un des concepts, le système l'ajoute au champ de saisie.

L'affichage des résultats. Une fois la requête analysée, le système récupère un certain nombre de documents correspondants. À chaque document, il associe un « marqueur sémantique » sur la base de la fréquence d'apparition de chacun des concepts métier de la requête. La corrélation de ces fréquences fournit un niveau de pertinence. La liste des documents est ensuite affichée par ordre décroissant de pertinence.

Que ce soit avec l'EyeTree ou avec l'OntoRequest, l'interface est divisée en deux avec à gauche le paradigme concerné et à droite, un volet pour afficher les documents. Ainsi l'utilisateur n'est pas perturbé en passant d'une carte à une autre.

5.2.6 Conclusion

La cartographie sémantique basée sur la sémantique du domaine permet une exploration de fonds documentaires qui répond aux besoins exprimés en termes de navigation, de vision globale et d'expression des besoins.

La modélisation des connaissances du domaine sous la forme d'une ontologie a permis non seulement de structurer le fonds documentaire via une annotation sémantique de son contenu mais aussi, de mettre en œuvre deux nouveaux paradigmes dédiés à la cartographie d'espaces informationnels structurés par une ontologie de domaine : le « EyeTree » et le « OntoRequest ». Ils ont été validés dans le cadre de la réalisation d'un système de cartographie sémantique de fonds documentaire MISTI (Moteur Internet Sémantique pour la Thermique Industrielle).

Le « EyeTree », navigateur graphique de type « fisheye polaire » (focus + context avec déformations linéaires) permet des interactions dédiées à l'exploration d'ontologies (parcours de sous éléments par rapport à un point fixe). Le « OntoRequest » permet de construire graphiquement une requête, basée sur la structuration ontologique des concepts métier du domaine, sans obliger les utilisateurs à maîtriser les connaissances du domaine. Il a l'avantage de permettre aux utilisateurs de généraliser et spécialiser leur requête intuitivement et sans surcharge d'informations.

La définition de ces deux outils n'a pas été immédiate. Elle a nécessité de nombreuses adaptations en accord avec les retours d'expérience des utilisateurs. Nous les avons donc obtenus suite à de nombreuses expérimentations et à plusieurs changements pouvant être de l'ordre de l'évolution (par exemple la modification des animations) ou de l'ordre du changement profond (comme le choix d'une toute nouvelle représentation).

Ces travaux ont permis de mettre en lumière l'importance de réaliser des cartographies basées sur la sémantique du domaine et sur les retours d'expérience. Cette conclusion est généralisable à d'autres domaines et il s'agit de l'objet même de la cartographie sémantique : cartographier les connaissances d'un domaine en exploitant sa sémantique.

5.3 Synthèse des mises en œuvres

Cette partie de ce document présente deux mises en œuvre de solutions de gestion de connaissances de l'équipe Condillac basée sur notre approche de la cartographie sémantique.

La première mise en œuvre correspond à un système de cartographie des compétences nommé Os Skill. Ce système est utilisé par plusieurs organisations dont la ville de Cran-Gevrier (400 utilisateurs) et le groupe « SADEC », un groupe d'expertise comptable (200 utilisateurs).

La seconde mise en œuvre est un système de cartographie pour l'exploration de fonds documentaires pour le « Groupement pour la Recherche sur les Échangeurs Thermiques » (GRETh), une émanation du CEA.

Le déploiement et la mise en œuvre de notre démarche et de nos réalisations dans le monde industriel valident cette étude. Ces deux cas témoignent de l'apport de notre approche de cartographie sémantique et les retours des utilisateurs ainsi que ceux des commanditaires de ces réalisations sont très positifs.

Partie 6 : **Conclusion & Perspectives**



« Connaître, ce n'est point démontrer, ni expliquer. C'est accéder à la vision. »

Antoine de Saint-Exupéry

6.1 Synthèse et conclusion

Cette étude a débuté avec un constat important : les organisations sont submergées par le nombre toujours croissant d'informations qu'il est nécessaire de maîtriser pour mener à bien leurs activités. En effet, étant obligées de faire face à la mondialisation des flux et des services, les organisations ont découvert de nouveaux besoins et elles ont développé de nouvelles activités : veille stratégique, gestion de l'innovation, capitalisation des savoirs, gestion du capital intellectuel et humain, intelligence économique, etc. Depuis, elles collectent et stockent toujours plus d'informations et leur espace informationnel ne cesse de grandir. Dans ce contexte mondialisé, l'avenir des organisations passe par la maîtrise de leur espace informationnel.

La maîtrise d'un espace est une problématique universelle qui préoccupe l'homme depuis son origine. La cartographie résulte de cette problématique ; elle permet de visualiser un espace pour mieux l'appréhender. Chaque carte joue alors le rôle de support externe de la pensée pour amplifier la cognition des personnes qui l'utilisent.

L'étude de la cartographie (géographique et thématique) et de la visualisation nous a appris que pour construire une carte d'un espace informationnel donné, il est nécessaire de mettre en œuvre des paradigmes de représentation, de visualisation et d'interaction. Cette « recette » est maintenant bien connue et maîtrisée. Cependant elle ne répond que partiellement à nos besoins. En effet, il subsiste deux problématiques à résoudre pour y parvenir :

- > Comment proposer une navigation basée sur la sémantique du domaine ?
- > Comment déterminer les paradigmes à mettre en œuvre pour fournir des cartes adaptées à l'utilisateur ?

Face à cette double problématique, nous avons défini une nouvelle cartographie capable de construire des cartes adaptées pour naviguer dans un espace informationnel selon la sémantique du domaine. Nous la nommons « **cartographie sémantique** ». Elle est présentée dans ce document au travers de nos différentes contributions : propositions, réalisations et mise en œuvre.

Nos propositions. Si différentes qu'elles puissent paraître, nos propositions constituent un tout cohérent, c'est-à-dire une **méthode** pour cartographier un espace informationnel selon nos besoins. Elle prend en compte la sémantique du domaine et exploite les retours d'expérience des utilisateurs.

Elle repose sur un nouveau **processus** qui assimile la cartographie à un ensemble d'opérations permettant de construire une carte d'un espace informationnel *brut* en passant successivement par un espace informationnel *structuré* puis un espace informationnel *représenté* et enfin un espace informationnel *visualisé*.

La mise en œuvre de ce processus a requis la création d'un **formalisme de description** dédié à la cartographie sémantique que nous avons nommé SNDF (Semantic Network Description Formalism). Il permet de décrire d'une part les connaissances du domaine en un ensemble d'entités et de relations entre ces entités, et d'autre part de décrire des propriétés de représentation pour ces entités et ces relations.

Nous avons aussi proposé des **critères pour la cartographie** d'espaces informationnels *structurés* par une ontologie de domaine (comme l'utilisation de techniques « focus+context » avec une géométrie euclidienne). La prise en compte de ces critères a abouti à la spécification

de deux nouveaux paradigmes pour explorer un espace informationnel *structuré* : l'EyeTree et le RadialTree.

La cartographie sémantique a aussi requis la définition d'un **langage** de haut niveau, nommé MDL (Map Description Language), pour construire une carte en décrivant sa cartographie sémantique.

Enfin, nous avons proposé une **architecture** de cartographie sémantique basée sur une extension du modèle MVC (Model View Controller). Elle permet d'intégrer la cartographie sémantique dans les solutions de gestion de connaissances à l'aide d'un ensemble d'agents logiciels.

Nos réalisations. Toutes les réalisations associées à cette étude ont été conçues dans le contexte particulier d'une Équipe de Recherche Technologie (ERT) associée à une société de valorisation leur conférant alors un caractère très industriel. La première réalisation présentée est Os Map, notre service web de cartographie sémantique. Ce service met directement en œuvre l'ensemble de nos propositions (méthode, processus, formalisme de description, paradigmes, langage et architecture). Ensuite, nous avons présenté SNCW. Il permet d'exploiter notre méthode pour construire graphiquement des modélisations de domaines. Enfin, nous avons décrit l'implémentation de l'EyeTree et du RadialTree à notre spécification de paradigmes pour l'exploration.

Nos mises en œuvre. Nous avons présenté deux mises en œuvre de solutions de gestion de connaissances de l'équipe Condillac basée sur notre approche de la cartographie sémantique. La première correspond à un système de cartographie des compétences nommé Os Skill. Il est utilisé par plusieurs organisations dont la ville de Cran-Gevrier (400 utilisateurs) et le groupe « SADEC », un groupe d'expertise comptable (200 utilisateurs). La seconde mise en œuvre est un système de cartographie de fonds documentaires pour le « Groupement pour la Recherche sur les Échangeurs Thermiques » (GRETh), une émanation du CEA.

L'objectif de ce travail était de lever un verrou technologique particulier : comment aider les utilisateurs à appréhender leur espace informationnel ? Notre réponse à cette problématique a donc été la définition de la cartographie sémantique. Sa mise en œuvre dans le monde des entreprises nous a permis de valider l'ensemble de l'étude mais surtout de s'assurer que ce verrou technologique identifié par le Ministère de la Recherche était bien levé.

Les retours des utilisateurs et de ceux qui ont commandité les différentes réalisations sont très positifs. Ils sont depuis convaincus que la cartographie sémantique constitue un atout pour leur démarche de gestion des connaissances ainsi que pour leurs différentes activités. Certains de ces commanditaires ont d'ailleurs témoigné dans la presse spécialisée de l'apport bénéfique de notre approche pour leurs activités.

Fort de ces résultats, il a été estimé que le verrou technologique de la cartographie sémantique était bien levé et, par conséquent que l'objectif de ce travail est atteint.

6.2 Perspectives

Nous identifions trois perspectives pour l'étude de la cartographie sémantique : perfectionner les outils, évaluer les paradigmes, adapter les cartes à chaque utilisateur et définir de nouveaux paradigmes.

Perfectionner les outils. Les outils présentés dans cette étude mettent en œuvre notre démarche de cartographie sémantique. Bien qu'ils permettent déjà de répondre à notre problématique, nous pouvons les perfectionner sur deux plans : individuel et général.

Sur le plan individuel, nous souhaitons rajouter de nouvelles fonctionnalités à nos outils. Par exemple, nous envisageons d'exporter les bases de connaissances de SNCW au format RDF et OWL pour être directement exploitables par des outils comme Protégé. Ou bien encore, nous développons actuellement la possibilité de construire des cartes dédiées à l'impression sur papier à partir de nos paradigmes EyeTree et RadialTree. Elles devront rester exploitables et efficaces malgré la perte d'interactivité due au papier. Ces nouvelles fonctionnalités ont pour objectif de toujours mieux répondre aux besoins des utilisateurs.

Sur le plan général, nous souhaitons faire évoluer nos outils pour exploiter d'avantage la sémantique exprimée par notre formalisme de description. Par exemple, nous voulons que la construction des représentations (les structures visuelles) des entités de l'espace informationnel *structuré* exploite la sémantique des relations qui les lient. Concrètement, nous envisageons que les entités liées à d'autres entités par des relations avec une sémantique d'héritage puissent « hériter » de leurs attributs graphiques. De même, nous pensons qu'il est possible d'exploiter graphiquement les relations qui ont une sémantique de composition. L'objectif est toujours d'exploiter au maximum la sémantique véhiculée par la modélisation du domaine.

Évaluer les paradigmes. Il ne faut pas omettre la dimension psychologique de la cartographie, c'est pourquoi nous estimons qu'il est nécessaire d'évaluer par des expérimentations l'impact de chaque paradigme sur les utilisateurs ainsi que leurs effets combinés. Concrètement, nous pensons que cette évaluation doit se traduire par l'analyse de l'activité de l'utilisateur c'est-à-dire par l'étude de l'activité de l'utilisateur lorsqu'il interagit avec les cartes (« analyse par observation du comportement des utilisateurs, puis, adaptation des techniques à l'objectif de conception et d'évaluation d'une interface » [Kovacs et al., 2004]). Cette approche cognitive complète dans une certaine mesure l'étude de satisfaction des utilisateurs (les retours d'expérience) par une approche statistique.

Ainsi en supposant une évaluation des paradigmes, de leur combinaison et de leurs impacts, leur choix pour constituer la cartographie peut alors être optimisé. En schématisant, on se plaît à imaginer une grille de choix à deux entrées pour déterminer les paradigmes à mettre en œuvre : la sémantique du domaine et l'évaluation des paradigmes.

Bien qu'une telle évaluation semble utopique dans l'état actuel de nos connaissances, nous estimons néanmoins que toute évaluation d'un paradigme permet de mieux en comprendre l'impact et par voie de conséquence de mieux l'utiliser.

Adapter les cartes à chaque utilisateur. Les utilisateurs interagissent avec leurs cartes (comme déplacer le point de vue ou changer de zoom) pour naviguer dans l'espace informationnel. Par soucis de construction de cartes adaptées, il nous paraît alors important d'étudier ces interactions. Ce qui va se traduire par la prise en compte de l'utilisateur (en tant qu'individu unique et non un type d'individu ou groupe) et par la mémorisation de son cheminement au sein des cartes.

L'objectif final est de réduire au maximum le temps entre l'utilisation d'une carte et son adaptation. Comme nous l'avons décrit dans notre approche, nous utilisons (à défaut de mieux) les retours d'expériences pour adapter le processus de cartographie. Mais ce cycle (utilisation, retours d'expérience, adaptation) est toujours trop long. Entre deux adaptations,

l'utilisateur est contraint d'utiliser une carte inadaptée (partiellement ou totalement) ce qui réduit nécessairement son efficacité.

Pour être efficace, une carte doit idéalement s'adapter en temps réel à son utilisateur. Cette perspective de recherche est une approche de cognition située où le couple utilisateur/carte est à mettre en symbiose.

Définir de nouveaux paradigmes. Nous pensons qu'il devient nécessaire de définir de nouveaux paradigmes pour augmenter les possibilités de cartographie et plus particulièrement pour changer les rapports entre l'utilisateur et l'espace informationnel. Dans l'état actuel des choses, les cartes constituent une sorte d'interface entre les informations et l'utilisateur. La carte constitue alors le monde virtuel et l'utilisateur le monde physique. Pour être encore plus efficace, il est nécessaire de réduire au maximum la frontière entre ces deux mondes.

Dans ce dessein nous identifions deux axes de recherche qui nous semblent très prometteurs : la « réalité virtuelle » et la « réalité augmentée ».

La réalité virtuelle permet d'immerger l'utilisateur dans l'espace informationnel. Elle correspond à la simulation par ordinateur d'un environnement visuel (souvent en 3D). Cependant bien que la révolution soit annoncée depuis plusieurs années, nous savons que pour le moment la réalité virtuelle ne répond pas complètement à ces attentes. Nous expliquons ce retard par deux faits : les matériels sont très onéreux et les utilisateurs supportent souvent mal l'immersion dans un monde virtuel. Ces derniers temps avec l'arrivée de matériel bon marché (casque 3D, gants, etc.) il ne reste plus qu'à définir de nouveaux paradigmes (représentation, visualisation et interaction) pour permettre aux utilisateurs de s'immerger efficacement dans un monde virtuel.

La réalité augmentée permet quant à elle d'immerger l'espace informationnel dans le monde réel. Un système de réalité augmentée permet d'intégrer des objets virtuels qu'on peut manipuler dans le monde réel. On peut par exemple imaginer des lunettes incrustant, automatiquement et en temps réel, des informations complémentaires sur ce que l'utilisateur voit (dans le monde réel). Tout comme avec la réalité virtuelle, les matériels sont onéreux et les paradigmes existants ne sont pas encore adaptés à notre problématique mais, la recherche dans ce domaine est très prolifique.

Nous estimons qu'à l'avenir ces deux axes de recherche offriront de nouveaux paradigmes pour construire des cartes permettant aux utilisateurs de toujours mieux maîtriser leurs espaces informationnels ; c'est l'objectif de la cartographie sémantique.

Partie 7 : **Bibliographie**



« Il faut que l'idée naisse de la vision comme l'étincelle du caillou. »

Charles-Ferdinand Ramuz

- [Ahlberg, 1996]
Ahlberg, C. (1996), '**Spotfire: an information exploration environment**', *SIGMOD Rec.*, 25, 4, p.25-29.
- [Ahlberg et al., 1992]
Ahlberg, C., Williamson, C. et Shneiderman, B. (1992), '**Dynamic queries for information exploration: an implementation and evaluation**', *Proceedings of the SIGCHI conference on Human factors in computing systems*, Monterey, California, United States, ACM Press
- [André, 1980]
André, A. (1980), *L'Expression Graphique : Cartes et diagrammes*, Masson.
- [Andrews, 1995]
Andrews, K. (1995), '**Case study: Visualising cyberspace: information visualisation in the Harmony Internet browser**', *Proceedings of the 1995 IEEE Symposium on Information Visualization*, Atlanta, Georgia, IEEE Computer Society
- [Andrews & Heidegger, 1998]
Andrews, K. et Heidegger, H. (1998), '**Information Slices: Visualising and Exploring Large Hierarchies using Cascading, Semi-Circular Discs**', *Late Breaking Hot Topic Paper, IEEE Symposium on Information Visualization (InfoVis'98)*, Research Triangle Park, North Carolina
- [Baader et al., 2003]
Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D. et Patel-Schneider, P.F. (eds.) (2003), *The description logic handbook: theory, implementation, and applications*, Cambridge University Press.
- [Bagot, 1996]
Bagot, J.-D. (1996), *Information, sensation et perception*, Paris, A. Colin.
- [Baker & Eick, 1995]
Baker, M.J. et Eick, S.G. (1995), '**Space-filling Software Visualization**', *Journal of Visual Languages and Computing*, 6, 2, p.119-133.
- [Baldonado et al., 2000]
Baldonado, M.Q.W., Woodruff, A. et Kuchinsky, A. (2000), '**Guidelines for using multiple views in information visualization**', *Proceedings of the working conference on Advanced visual interfaces*, Palermo, Italy, ACM Press
- [Barlow & Padraic, 2001]
Barlow, S.T. et Padraic, N. (2001), '**A Comparison of 2-D Visualizations of Hierarchies**', *Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01)*, IEEE Computer Society
- [Barr & Feigenbaum, 1982]
Barr, A. et Feigenbaum, E.A. (1982), *The handbook of Artificial Intelligence (Vol.1)*, Addison Wesley, New York.
- [Bechhofer et al., 2001]
Bechhofer, S., Horrocks, I., Goble, C. et Stevens, R. (2001), '**OilEd: a Reason-able Ontology Editor for the Semantic Web**', *Proceedings of KI2001, Joint German/Austrian conference on Artificial Intelligence*, Vienna, Springer-Verlag, p.396-408.
- [Bederson, 2000]
Bederson, B.B. (2000), '**Fisheye menus**', *Proceedings of the 13th annual ACM symposium on User interface software and technology*, San Diego, California, United States, ACM Press
- [Bertin, 1977]
Bertin, J. (1977), *La graphique et le traitement graphique de l'information*, Paris, Flammarion.
- [Bertin, 1999 (1re éd. 1969)]
Bertin, J. (1999 (1re éd. 1969)), *Sémiologie graphique : les diagrammes, les réseaux, les cartes* (3ème Edition), Paris, Editions de l'Ecole des Hautes Etudes en Sciences.
- [Bonin, 1997]
Bonin, S. (1997), '**Le développement de la graphique de 1967 à 1997**', *30 ans de sémiologie graphique*, Paris
- [Börner et al., 2005]
Börner, K., Dall'Asta, L., Ke, W. et Vespignani, A. (2005), '**Studying the emerging global brain: Analyzing and visualizing the impact of co-authorship teams**', *Complexity*, 10, 4, p.57-67.
- [Burkhard, 2004]
Burkhard, R.A. (2004), '**Learning from Architects: The Difference between Knowledge Visualization and Information Visualization**', *Proceedings of the Information Visualisation, Eighth International Conference on (IV'04) - Volume 00*, IEEE Computer Society
- [Burkhard, 2005]

-
- Burkhard, R.A. (2005), '**Towards a Framework and a Model for Knowledge Visualization: Synergies Between Information and Knowledge Visualization**', *Knowledge and information visualization : searching for synergies*, 3426, p.238-255.
- [Card et al., 2004]
Card, S.K., Hong, L., Mackinlay, J.D. et Chi, E.H. (2004), '**3Book: a scalable 3D virtual book**', *CHI '04 extended abstracts on Human factors in computing systems*, Vienna, Austria, ACM Press
- [Card et al., 1999a]
Card, S.K., Mackinlay, J.D. et Shneiderman, B. (1999a), '**Data mapping: document visualization**', *Readings in information visualization: using vision to think*, Morgan Kaufmann Publishers Inc., p.409-410.
- [Card et al., 1999b]
Card, S.K., Mackinlay, J.D. et Shneiderman, B. (1999b), '**Information visualization**', *Readings in information visualization: using vision to think*, Morgan Kaufmann Publishers Inc., p.1-34.
- [Card et al., 1999c]
Card, S.K., Mackinlay, J.D. et Shneiderman, B. (1999c), '**Overview + detail**', *Readings in information visualization: using vision to think*, Morgan Kaufmann Publishers Inc., p.285-286.
- [Card et al., 1999d]
Card, S.K., Mackinlay, J.D. et Shneiderman, B. (1999d), *Readings in information visualization : using vision to think*, San Francisco, Calif., Morgan Kaufmann Publishers,.
- [Card et al., 1994]
Card, S.K., Pirolli, P. et Mackinlay, J.D. (1994), '**The cost-of-knowledge characteristic function: display evaluation for direct-walk dynamic information visualizations**', *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*, Boston, Massachusetts, United States, ACM Press
- [Card et al., 1996]
Card, S.K., Robertson, G.G. et York, W. (1996), '**The WebBook and the Web Forager: an information workspace for the World-Wide Web**', *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground*, Vancouver, British Columbia, Canada, ACM Press
- [Chabris & Kosslyn, 2005]
Chabris, C.F. et Kosslyn, S.M. (2005), '**Representational correspondence as a basic principle of diagram design**', *Knowledge and information visualization : searching for synergies*, 3426.
- [Chambers et al., 1983]
Chambers, J.M., Cleveland, W.S. et Tukey, P.A. (1983), *Graphical Methods for Data Analysis*, New York, Chapman and Hall.
- [Chang et al., 2002]
Chang, D., Dooley, L. et Tuovinen, J.E. (2002), '**Gestalt theory in visual screen design: a new look at an old subject**', *Proceedings of the Seventh world conference on computers in education conference on Computers in education: Australian topics - Volume 8*, Copenhagen, Denmark, Australian Computer Society, Inc.
- [Chernoff, 1973]
Chernoff, H. (1973), '**The Use of Faces to Represent Points in k-Dimensional Space Graphically**', *Journal of the American Statistical Association*, 68, 342, p.361\ -368.
- [Chi & Riedl, 1998]
Chi, E.H.-h. et Riedl, J.T. (1998), '**An Operator Interaction Framework for Visualization Systems**', p.63--70.
- [Chi, 2000]
Chi, E.H. (2000), '**A Taxonomy of Visualization Techniques Using the Data State Reference Model**', *Proceedings of the IEEE Symposium on Information Visualization 2000*, IEEE Computer Society, p.69.
- [Chuah & Roth, 1996]
Chuah, M.C. et Roth, S.F. (1996), '**On the semantics of interactive visualizations**', *Proceedings of the 1996 IEEE Symposium on Information Visualization (INFOVIS '96)*, IEEE Computer Society
- [Corman & Ingargiola, 2005]
Corman, V. et Ingargiola, E. (2005), '**Guide pratique :intelligence économique et PME.**' in MEDEF (ed.), Paris
- [Cubaud et al., 1998]
Cubaud, P., Thiria, C. et Topol, A. (1998), '**Experimenting a 3D interface for the access to a digital library**', *Proceedings of the third ACM conference on Digital libraries*, Pittsburgh, Pennsylvania, United States, ACM Press
- [Cubaud & Topol, 2001]

- Cubaud, P. et Topol, A. (2001), '**A VRML-based user interface for an online digitalized antiquarian collection**', *Proceedings of the sixth international conference on 3D Web technology*, Paderbon, Germany, ACM Press
- [Dansereau, 2005]
Dansereau, D.F. (2005), '**Node-Link Mapping Principles for Visualizing Knowledge and Information.**' *Knowledge and information visualization : searching for synergies*, 3426, p.61-81.
- [Dejoux, 2000]
Dejoux, C. (2000), '**Le point sur les approches COMPÉTENCE : Pour une démarche globale et agrégée**', *IFSAM*, HEC Montréal
- [Descartes, 1637]
Descartes, R. (1637), *La Dioptrique*.
- [Dieng-Kuntz et al., 2005]
Dieng-Kuntz, R., Corby, O., Gandon, F., Giboin, A., Golebiowska, J., Matta, N. et Ribière, M. (2005), '**Méthodes et outils pour la gestion des connaissances : Une approche pluridisciplinaire du Knowledge Management**' (Dunod Edition).
- [dos Santos & Brodlie, 2004]
dos Santos, S. et Brodlie, K. (2004), '**Gaining understanding of multivariate and multidimensional data through visualization.**' *Computers & Graphics*, 28, 3, p.311-325.
- [Eick et al., 1992]
Eick, S.G., Steffen, J.L. et Jr., E.E.S. (1992), '**Seesoft-A Tool For Visualizing Line Oriented Software Statistics.**' *IEEE Trans. Software Eng.*, 18, 11, p.957-968.
- [Furnas, 1981]
Furnas, G.W. (1981), '**The FISHEYE View: A New Look at Structured Files**'
- [Furnas, 1986]
Furnas, G.W. (1986), '**Generalized fisheye views**', *Proceedings of the SIGCHI conference on Human factors in computing systems*, Boston, Massachusetts, United States, ACM Press
- [Furnas & Bederson, 1995]
Furnas, G.W. et Bederson, B.B. (1995), '**Space-scale diagrams: understanding multiscale interfaces**', *Proceedings of the SIGCHI conference on Human factors in computing systems*, Denver, Colorado, United States, ACM Press/Addison-Wesley Publishing Co.
- [Gardarin, 2002]
Gardarin, G. (2002), *XML : Des bases de données aux services Web*.
- [Ghoniem et al., 2004]
Ghoniem, M., Fekete, J.-D. et Castagliola, P. (2004), '**A Comparison of the Readability of Graphs Using Node-Link and Matrix-Based Representations**', *Proceedings of the IEEE Symposium on Information Visualization (INFOVIS'04) - Volume 00*, IEEE Computer Society
- [Gordon, 2000]
Gordon, J.L. (2000), '**Creating knowledge maps by exploiting dependent relationships.**' *Knowledge Based Systems*, 13, 2-3, p.71-79.
- [Gregory, 2000]
Gregory, R.L. (2000), *L'œil et le cerveau : La psychologie de la vision* (De Boeck Université Edition).
- [Gruber, 1993]
Gruber, T.R. (1993), '**A translation approach to portable ontology specifications**', *Knowledge Acquisition*, 5, 2, p.199-220.
- [Guarino & Giaretta, 1995]
Guarino, N. et Giaretta, P. (1995), '**Ontologies and Knowledge Bases: Towards a Terminological Clarification**', *Towards Very Large Knowledge Bases*, IOS Press, Amsterdam, p.25-32.
- [Guillaume, 1937]
Guillaume, P. (1937), *La psychologie de la forme*, Paris, Flammarion.
- [Hascoët, 2001]
Hascoët, M. (2001), '**Interaction and Visualization Supporting Web Browsing Patterns**', *IV'2001 - Information Visualisation*, London, IEEE Computer Society, p.413-419.
- [Hascoët & Beaudouin-Lafon, 2001]
Hascoët, M. et Beaudouin-Lafon, M. (2001), '**Visualisation Interactive d'Information**', *I3: Information, Interaction, Intelligence*, Vol. 1, n° 1, p.77-108.
- [Heer & Boyd, 2005]
Heer, J. et Boyd, D. (2005), '**Vizster: Visualizing Online Social Networks**', *Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization*, IEEE Computer Society
- [Herman et al., 2000]

-
- Herman, I., Melançon, G. et Marshall, M.S. (2000), '**Graph Visualization and Navigation in Information Visualization: A Survey**', *IEEE Transactions on Visualization and Computer Graphics*, 6, 1, p.24-43.
- [Hernandez, 2005]
Hernandez, N. (2005), '**Ontologies pour l'aide à l'exploration d'une collection de documents**', *Ingénierie des Systèmes d'Information* Hermès Sciences p.11-31.
- [Hernandez & Aussenac-Gilles, 2004]
Hernandez, N. et Aussenac-Gilles, N. (2004), '**OntoExplo : Ontologies pour l'aide à une activité de veille ou d'exploration d'un domaine**', *VIème Journées de l'innovation*, Foix
- [Hughes et al., 2004]
Hughes, T., Hyun, Y. et Liberles, D.A. (2004), '**Visualising very large phylogenetic trees in three dimensional hyperbolic space**', *BMC Bioinformatics*, 5, 1, p.48.
- [Inselberg & Dimsdale, 1990]
Inselberg, A. et Dimsdale, B. (1990), '**Parallel coordinates: a tool for visualizing multi-dimensional geometry**', *Proceedings of the 1st conference on Visualization '90*, San Francisco, California, IEEE Computer Society Press
- [Jacquemin et al., 2005]
Jacquemin, C., Folch, H., Garcia, K. et Nugier, S. (2005), '**Visualisation interactive d'espaces documentaires**.' *Revue I3*, 5, 1.
- [Jaeschke et al., 2005]
Jaeschke, G., Leissler, M. et Hemmje, M. (2005), '**Modeling Interactive, 3-Dimensional Information Visualizations Supporting Information Seeking Behaviors**.' *Knowledge and information visualization : searching for synergies*, 3426, p.119-135.
- [Jerding & Stasko, 1995]
Jerding, D.F. et Stasko, J.T. (1995), '**The information mural: a technique for displaying and navigating large information spaces**', *Proceedings of the 1995 IEEE Symposium on Information Visualization*, Atlanta, Georgia, IEEE Computer Society
- [Johnson & Shneiderman, 1991]
Johnson, B. et Shneiderman, B. (1991), '**Tree-Maps: a space-filling approach to the visualization of hierarchical information structures**', *Proceedings of the 2nd International IEEE Visualization Conference*, San Diego, p.284-291.
- [Kadima & Monfort, 2003]
Kadima, H. et Monfort, V. (2003), *Les Web services - Techniques, démarches et outils : XML, WSDL, SOAP, UDDI, Rosetta, UML*.
- [Keim, 2002]
Keim, D.A. (2002), '**Information Visualization and Visual Data Mining**', *IEEE Transactions on Visualization and Computer Graphics*, 8, 1, p.1-8.
- [Keller & Tergan, 2005]
Keller, T. et Tergan, S.-O. (2005), '**Visualizing Knowledge and Information: An Introduction**', *Knowledge and information visualization : searching for synergies*, 3426, p.1-23.
- [Kiryakov et al., 2003]
Kiryakov, A., Popov, B., Ognyanoff, D., Manov, D., Kirilov, A. et Goranov, M. (2003), '**Semantic Annotation, Indexing, and Retrieval**', in Mylopoulos, J. (ed.), *International Semantic Web Conference*, Springer, p.484-499.
- [Kleiberg et al., 2001]
Kleiberg, E., van de Wetering, H. et van Wijk, J.J. (2001), '**Botanical Visualization of Huge Hierarchies**', *Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01)*, IEEE Computer Society
- [Kobsa, 2004]
Kobsa, A. (2004), '**User Experiments with Tree Visualization Systems**'.
- [Kovacs et al., 2004]
Kovacs, B., Gaunet, F. et Briffault, X. (2004), *Les techniques d'analyse de l'activité pour l'IHM*, Hermès - Lavoisier.
- [Krasner & Pope, 1988]
Krasner, G.E. et Pope, S.T. (1988), '**A cookbook for using the model-view controller user interface paradigm in Smalltalk-80**', *J. Object Oriented Program.*, 1, 3, p.26-49.
- [Lamping et al., 1995]
Lamping, J., Rao, R. et Pirolli, P. (1995), '**A focus+context technique based on hyperbolic geometry for visualizing large hierarchies**', *Proceedings of the SIGCHI conference on Human factors in computing systems*, Denver, Colorado, United States, ACM Press/Addison-Wesley Publishing Co.
- [Latour, 1985]

- Latour, B. (1985), '**Les « Vues » de l'Esprit**', *Culture Technique*, 14, p.4-30.
- [Lee et al., 2006]
Lee, B., Parr, C.S., Plaisant, C., Bederson, B.B., Veksler, V.D., Gray, W.D. et Kotfila, C. (2006), '**TreePlus: Interactive Exploration of Networks with Enhanced Tree Layouts**', *IEEE Transactions on Visualization and Computer Graphics*, Los Alamitos, CA, USA, IEEE Computer Society, p.1414-1426.
- [Leung & Apperley, 1994]
Leung, Y.K. et Apperley, M.D. (1994), '**A review and taxonomy of distortion-oriented presentation techniques**', *ACM Trans. Comput.-Hum. Interact.*, 1, 2, p.126-160.
- [Leuski & Allan, 2000]
Leuski, A. et Allan, J. (2000), '**Lighthouse: Showing the Way to Relevant Information**', *Proceedings of the IEEE Symposium on Information Visualization 2000*, IEEE Computer Society
- [Lohse et al., 1994]
Lohse, G.L., Biolsi, K., Walker, N. et Rueter, H.H. (1994), '**A classification of visual representations**', *Commun. ACM*, 37, 12, p.36-49.
- [Mackinlay et al., 1991]
Mackinlay, J.D., Robertson, G.G. et Card, S.K. (1991), '**The perspective wall: detail and context smoothly integrated**', *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*, New Orleans, Louisiana, United States, ACM Press
- [Mane & Borner, 2004]
Mane, K.K. et Borner, K. (2004), '**Mapping topics and topic bursts in PNAS**', *Proceeding of the National Academy of Sciences of the United States* (101 Edition), p.5287-5290.
- [Marey, 1878]
Marey, É.-J. (1878), *La méthode graphique dans les sciences expérimentales et principalement en physiologie et en médecine*.
- [Minsky, 1975]
Minsky, M. (1975), '**A Framework for Representing Knowledge**', *The Psychology of Computer Vision*, McGraw Hill, New York, p.211-277.
- [Mothe et al., 2003]
Mothe, J., Chrismont, C., Dousset, B. et Alaux, J. (2003), '**DocCube: Multi-Dimensional Visualisation and Exploration of Large Document Sets**', *Journal of the American Society for Information Science and Technology, JASIST, Special topic section: web retrieval and mining*, 7, 54, p.650-659.
- [Munzner & Burchard, 1995]
Munzner, T. et Burchard, P. (1995), '**Visualizing the structure of the World Wide Web in 3D hyperbolic space**', *Proceedings of the first symposium on Virtual reality modeling language*, San Diego, California, United States, ACM Press
- [Napoli, 1997]
Napoli, A. (1997), '**Une introduction aux logiques de descriptions**'
- [Nation, 1998]
Nation, D.A. (1998), '**WebTOC: a tool to visualize and quantify Web sites using a hierarchical table of contents browser**', *CHI 98 conference summary on Human factors in computing systems*, Los Angeles, California, United States, ACM Press
- [Neumann et al., 2005]
Neumann, A., Gräber, W. et Tergan, S.-O. (2005), '**ParIS - Visualizing Ideas and Information in a Resource-Based Learning Scenario**', *Knowledge and information visualization : searching for synergies*, 3426, p.256-281.
- [Nilsson, 1980]
Nilsson, N.J. (1980), *Principles of artificial intelligence*, Morgan Kaufmann Publishers Inc.
- [Nishizeki & Rahman, 2004]
Nishizeki, T. et Rahman, M.S. (2004), *Planar Graph Drawing*, World Scientific Publishing Company.
- [Norman, 1993]
Norman, D.A. (1993), *Things that make us smart: defending human attributes in the age of the machine*, Addison-Wesley Longman Publishing Co., Inc.
- [North & Schneiderman, 1997]
North, C. et Schneiderman, B. (1997), '**A Taxonomy of Multiple Window Coordinations**'
- [Novak, 2001]
Novak, J.D. (2001), '**The Theory Underlying Concept Maps and How to Construct Them**'
- [Olsen et al., 1993]
Olsen, K.A., Korfhage, R.R., Sochats, K.M., Spring, M.B. et Williams, J.G. (1993), '**Visualization of a document collection: the vibe system**', *Inf. Process. Manage.*, 29, 1, p.69-81.

-
- [Pirolli et al., 2003]
Pirolli, P., Card, S.K. et van Der Wege, M.M. (2003), '**The effects of information scent on visual search in the hyperbolic tree browser**', *ACM Trans. Comput.-Hum. Interact.*, 10, 1, p.20-53.
- [Plaisant, 2004]
Plaisant, C. (2004), '**The challenge of information visualization evaluation**', *Proceedings of the working conference on Advanced visual interfaces*, Gallipoli, Italy, ACM Press
- [Plaisant et al., 1995]
Plaisant, C., Carr, D. et Shneiderman, B. (1995), '**Image-Browser Taxonomy and Guidelines for Designers**', *IEEE Softw.*, 12, 2, p.21-32.
- [Poidevin, 1999]
Poidevin, D. (1999), *La carte, moyen d'action* (Ellipses Marketing Edition).
- [Purchase, 1997]
Purchase, H.C. (1997), '**Which Aesthetic has the Greatest Effect on Human Understanding?**' *Proceedings of the 5th International Symposium on Graph Drawing*, Springer-Verlag
- [Rao & Card, 1994]
Rao, R. et Card, S.K. (1994), '**The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus+Context Visualization for Tabular Information**', in Beth, A., Susan, D. et Judith, S.O. (eds.), ACM Press, p.318-322.
- [Rekimoto & Green, 1993]
Rekimoto, J. et Green, M. (1993), '**The information cube: Using transparency in 3d information visualization**'
- [Rennison, 1994]
Rennison, E. (1994), '**Galaxy of news: an approach to visualizing and understanding expansive news landscapes**', *Proceedings of the 7th annual ACM symposium on User interface software and technology*, Marina del Rey, California, United States, ACM Press
- [Robertson et al., 1998]
Robertson, G.G., Czerwinski, M., Larson, K., Robbins, D.C., Thiel, D. et van Dantzich, M. (1998), '**Data mountain: using spatial memory for document management**', *Proceedings of the 11th annual ACM symposium on User interface software and technology*, San Francisco, California, United States, ACM Press
- [Robertson & Mackinlay, 1993]
Robertson, G.G. et Mackinlay, J.D. (1993), '**The document lens**', *Proceedings of the 6th annual ACM symposium on User interface software and technology*, Atlanta, Georgia, United States, ACM Press
- [Robertson et al., 1991]
Robertson, G.G., Mackinlay, J.D. et Card, S.K. (1991), '**Cone Trees: animated 3D visualizations of hierarchical information**', *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*, New Orleans, Louisiana, United States, ACM Press
- [Roche, 2003]
Roche, C. (2003), '**Ontology : a survey**', *8th IFAC Symposium on Automated Systems Based on Human Skill and Knowledge*, Göteborg, Sweden
- [Roche, 2005]
Roche, C. (2005), '**Terminologie & Ontologie**', *Langages*, 157, p.48-62.
- [Roche et al., 2005]
Roche, C., Foveau, C.-E. et Reguigui, S. (2005), '**La démarche ontologique pour la gestion des compétences et des connaissances**', *Revue des Nouvelles Technologies de l'Information*, RNTI-E-3, p.359-364.
- [Roche & Million-Rousseau, 2003]
Roche, C. et Million-Rousseau, C. (2003), '**Construction de Terminologies Métier : L'Importance du Modèle Ontologique**', *Journées francophones d'Extraction et de Gestion des Connaissances (EGC 2003)*
- [Rouleau, 2000]
Rouleau, B. (2000), *Méthode de la cartographie* (CNRS Editions Edition).
- [Sarkar & Brown, 1992]
Sarkar, M. et Brown, M.H. (1992), '**Graphical fisheye views of graphs**', *Proceedings of the SIGCHI conference on Human factors in computing systems*, Monterey, California, United States, ACM Press
- [Scaife & Rogers, 1996]
Scaife, M. et Rogers, Y. (1996), '**External cognition: how do graphical representations work?**' *Int. J. Hum.-Comput. Stud.*, 45, 2, p.185-213.
- [Shneiderman, 1992]
Shneiderman, B. (1992), '**Tree visualization with tree-maps: 2-d space-filling approach**', *ACM Trans. Graph.*, 11, 1, p.92-99.

- [Shneiderman, 1996]
Shneiderman, B. (1996), '**The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations**', *IEEE Symposium on Visual Languages*, IEEE Computer Society
- [Sowa, 2000]
Sowa, J.F. (2000), *Knowledge representation : logical, philosophical, and computational foundations*, Pacific Grove, Calif., Brooks/Cole.
- [Spence & Apperley, 1999]
Spence, R. et Apperley, M. (1999), '**Data base navigation: an office environment for the professional**', *Readings in information visualization: using vision to think*, Morgan Kaufmann Publishers Inc., p.333-340.
- [Stasko & Zhang, 2000]
Stasko, J. et Zhang, E. (2000), '**Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations**', *Proceedings of the IEEE Symposium on Information Visualization 2000*, IEEE Computer Society
- [Stone et al., 1994]
Stone, M.C., Fishkin, K. et Bier, E.A. (1994), '**The movable filter as a user interface tool**', *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*, Boston, Massachusetts, United States, ACM Press
- [Tractinsky, 1997]
Tractinsky, N. (1997), '**Aesthetics and apparent usability: empirically assessing cultural and methodological issues**', *SIGCHI conference on Human factors in computing systems*, Atlanta, Georgia, United States, ACM Press, p.115-122.
- [Tufte, 1986]
Tufte, E.R. (1986), *The visual display of quantitative information*, Graphics Press.
- [Tweedie, 1997]
Tweedie, L. (1997), '**Characterizing interactive externalizations**', *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, p.375--382.
- [Uschold & Grüninger, 1996]
Uschold, M. et Grüninger, M. (1996), '**Ontologies: principles, methods, and applications**', *Knowledge Engineering Review*, 11, 2, p.93-155.
- [van Ham & van Wijk, 2003]
van Ham, F. et van Wijk, J.J. (2003), '**Beamtrees: compact visualization of large hierarchies**', *Information Visualization*, 2, 1, p.31-39.
- [van Wijk & Nuij, 2003]
van Wijk, J.J. et Nuij, W.A.A. (2003), '**Smooth and efficient zooming and panning**', *infovis*, Los Alamitos, CA, USA, IEEE Computer Society, p.3.
- [van Wijk & van de Wetering, 1999]
van Wijk, J.J. et van de Wetering, H. (1999), '**Cushion Treemaps: Visualization of Hierarchical Information**', *Proceedings of the 1999 IEEE Symposium on Information Visualization*, IEEE Computer Society
- [van Wijk et al., 2003]
van Wijk, J.J., van Ham, F. et van de Wetering, H. (2003), '**Rendering hierarchical data**', *Commun. ACM*, 46, 9, p.257-263.
- [Ware, 2000]
Ware, C. (2000), *Information visualization: perception for design*, Morgan Kaufmann Publishers Inc.
- [Ware, 2005]
Ware, C. (2005), '**Visual Queries: The Foundation of Visual Thinking**', *Knowledge and information visualization : searching for synergies*, 3426, p.27-35.
- [Ware et al., 2002]
Ware, C., Purchase, H., Colpoys, L. et McGill, M. (2002), '**Cognitive measurements of graph aesthetics**', *Information Visualization*, 1, 2, p.103-110.
- [Wattenberg & Fisher, 2003]
Wattenberg, M. et Fisher, D. (2003), '**A Model of Multi-Scale Perceptual Organization in Information Graphics**', *Infovis*, 00, p.4.
- [Wise et al., 1995]
Wise, J.A., Thomas, J.J., Pennock, K., Lantrip, D., Pottier, M., Schur, A. et Crow, V. (1995), '**Visualizing the non-visual: spatial analysis and interaction with information from text documents**', *Proceedings of the 1995 IEEE Symposium on Information Visualization*, Atlanta, Georgia, IEEE Computer Society
- [Wiss & Carr, 1998]
Wiss, U. et Carr, D. (1998), '**A Cognitive Classification Framework for 3-Dimensional Information Visualization**'

[Woolridge, 2001]

Woolridge, M. (2001), *Introduction to Multiagent Systems*, John Wiley & Sons, Inc.

[Zhang, 1997]

Zhang, J. (1997), 'The Nature of External Representations in Problem Solving', *Cognitive Science*, 21, 2, p.179-216.

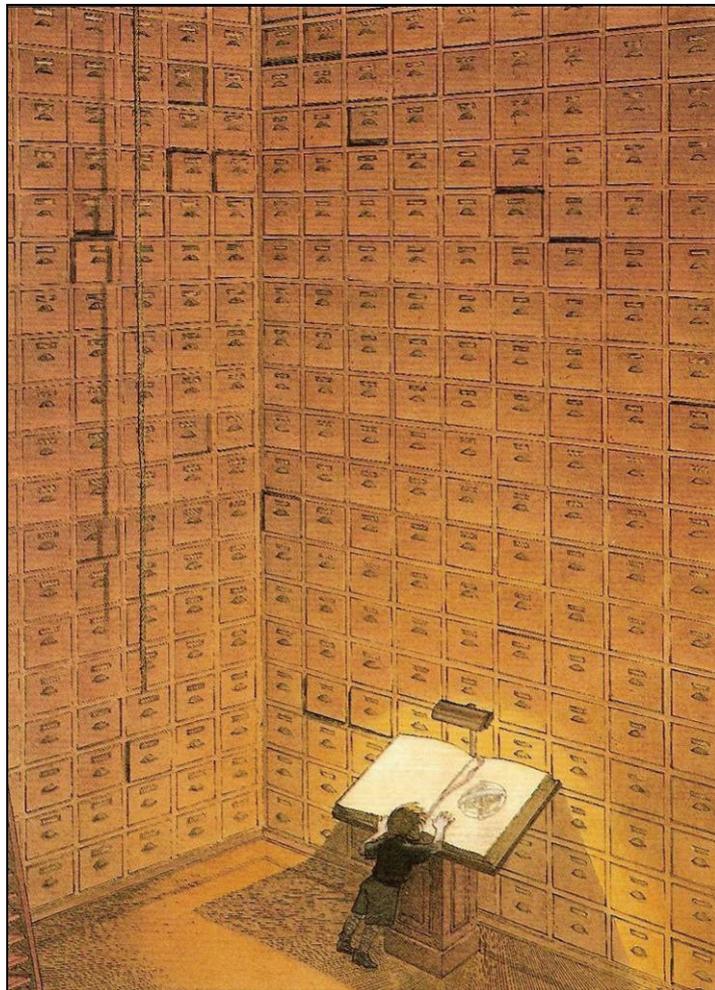
[Zhang & Norman, 1994]

Zhang, J. et Norman, D.A. (1994), 'Representations in Distributed Cognitive Tasks', *Cognitive Science*, 18, 1, p.87-122.

[Zlatoff et al., 2004]

Zlatoff, N., Tellez, B. et Baskurt, A. (2004), 'Vision Gestalt et connaissances : une approche générique à l'interprétation d'images', *CORESAs 2004*, Lille

Partie 8 : Annexes



« Un bon croquis vaut mieux qu'un long discours. »

Napoléon Bonaparte

Rappel du plan

8.1 Lexique	239
8.2 Index.....	241
8.3 La perception visuelle	243
8.4 MDL.....	245
8.4.1 DTD de MDL.....	245
8.4.2 Exemples de cartes avec MDL	248
8.5 Description WSDL du service Os Map.....	254
8.6 Tables	256
8.6.1 Table des figures	256
8.6.2 Table des tableaux	259
8.7 Crédits images	260
8.7.1 Les cités obscures	260
8.7.2 Circle Limite IV.....	260

8.1 Lexique

Carte

La carte est une représentation graphique conventionnelle en positions relatives, de phénomènes concrets ou abstraits, localisables dans un espace.

Cartographe

Le cartographe est un opérateur capable de maîtriser les concepts et les taxons de la discipline qu'il se charge d'illustrer sous forme de carte.

Cartographie

La cartographie est l'ensemble des opérations qui interviennent en vue de l'élaboration d'une carte et de son utilisation.

Cartographie de données abstraites

La cartographie de données abstraites est la cartographie d'un ensemble de données abstraites.

Cartographie géographique

La cartographie est l'ensemble des études et des opérations scientifiques, artistiques et techniques, intervenant à partir des résultats d'opérations directes ou de l'exploitation d'une documentation, en vue de l'élaboration et de l'établissement de cartes, plans et autres modes d'expression, ainsi que dans leur utilisation.

Cartographie sémantique

La cartographie sémantique est la cartographie de l'espace informationnel d'une organisation basée sur sa sémantique (comprise comme une conceptualisation du domaine).

Cognition

Fonction complexe multiple regroupant l'ensemble des activités mentales (pensée, perception, action, volonté, mémorisation, rappel, apprentissage) impliquées dans la relation de l'être humain avec son environnement et qui lui permettent d'acquérir et de manipuler des connaissances (associations, rétroaction, traitement de l'information, résolution de problèmes, prise de décision etc.).

Cognition externe

Utilisation de la perception de supports externes pour amplifier les mécanismes cognitifs. La cognition externe est un mécanisme du raisonnement distribué.

Graphique (la)

La construction d'images à partir d'une grammaire qui s'appuie sur les lois de la perception visuelle.

Modélisation

Représentation explicite de la sémantique d'un domaine par un ensemble des concepts permettant de décrire les objets du monde. Ces concepts sont décrits selon une théorie donnée.

Sémantique (d'un domaine)

Ensemble des concepts permettant d'appréhender les objets du monde (le domaine).

Visualisation

La visualisation est l'utilisation de représentations visuelles de données pour amplifier la cognition.

Visualisation de connaissances

La visualisation de connaissances est l'utilisation de représentations visuelles de connaissances à des fins de création et de partage.

Visualisation d'informations

La visualisation d'informations est l'utilisation de représentations visuelles de données abstraites pour amplifier la cognition.

8.2 Index

A

Agent logiciel..... 12, 147, 165
 Aide externe.....31
 Aide-mémoire..... 4, 25, 32
 Amplification..... 10, 11, 82, 132, 135, 189, 190
 Annotation sémantique.....9, 126, 212, 219
 Arbre de cônes.....69, 129, 130, 169
 Architecture
 agents (à)..... 11, 13, 147, 148, 163, 171, 182
 cartographie sémantique (de) 147, 148, 153, 154, 156,
 223

B

Base de connaissances9, 13, 125, 126, 127, 128, 129, 132,
 172, 181, 182, 183, 185, 186, 188, 189
 Beamtrees..... 74, 128
 Bifocales.....82, 83, 84, 86
 Brushing.....89

C

Carte..... 4, 29, 203, 239
 conceptuelle..... 50, 175
 géographique.....44
 heuristique..... 175
 sémantique.....97
 stratégique.....206
 thématique.....28
 Cartographe.....4, 29, 87, 140, 152, 153, 163, 167, 239
 Cartographie
 antique..... 4, 25
 données abstraites (de)..... 41, 239
 géographique..... 25, 30, 44
 sémantique.....2, 4, 5, 6, 8, 9, 12, 13, 14, 15, 18, 21, 25,
 43, 93, 94, 95, 97, 98, 99, 100, 103, 105, 109, 111,
 124, 125, 126, 127, 131, 135, 137, 138, 144, 146,
 147, 148, 149, 151, 153, 155, 156, 159, 160, 162,
 163, 171, 172, 178, 188, 195, 199, 200, 201, 208,
 210, 212, 214, 219, 220, 222, 223, 225, 239, 248
 thématique..... 4, 27, 28
 topographique..... 4, 26, 27
 Chernoff (figures de)..... 59, 62
 Cités obscures (les).....260
 Classification..... 5, 55, 107, 138, 203, 212
 Cognition.....239
 Cognition externe.....30, 31, 32, 39, 239
 Contrôle de point de vue.....53, 79, 80, 89
 Coordonnées parallèles..... 59, 60, 61, 67

D

Description de représentation..... 122
 Détails à la demande.....89

Diagrammes en étoile.....59, 61
 Direct walk.....88
 Disques d'informations.....75, 128
 Domaine de signification..... 34

E

Environnement de modélisation..... 172, 176, 178, 182
 Équipe de recherche technologique (ERT)12, 14, 159, 199
 Espace
 carte (de la).....45, 57, 68, 72, 79, 80, 86, 96, 107, 144,
 169
 informationnel.....3, 7, 19, 54, 97, 99, 125
 brut.....5, 8, 93, 103, 106, 107, 108, 151, 152, 156,
 222
 représenté..... 5, 7, 8, 53, 55, 78, 79, 85, 88, 93, 99,
 102, 103, 106, 107, 108, 121, 132, 134, 138,
 141, 156, 167, 189, 222
 structuré..... 5, 6, 7, 8, 9, 11, 93, 101, 102, 103, 106,
 107, 108, 109, 110, 111, 116, 120, 121, 123,
 124, 125, 132, 135, 137, 138, 139, 141, 142,
 143, 146, 152, 156, 164, 166, 167, 168, 178,
 189, 194, 201, 222, 223
 visualisé..... 5, 7, 8, 53, 55, 78, 89, 93, 99, 102, 103,
 106, 107, 108, 132, 141, 156, 222
 représentation de)..... 10, 79, 106, 110, 120, 192, 194
 Évaluation..... 14, 127, 205, 224
 Exploration.....2, 9, 13, 50, 98, 125, 126, 135, 159, 162, 163,
 171, 189, 194, 195, 200, 212, 214, 217, 218, 219, 220,
 223
 Extralinguistique.....211, 212

F

Fisheye..... 11, 82, 85, 86, 87, 130, 135, 189, 192, 193, 219
 Focus + context..... 6, 93, 156, 222
 Fonction d'amplification..... 190
 Formalisme
 de description2, 6, 8, 9, 11, 13, 93, 109, 111, 116, 118,
 123, 124, 138, 140, 146, 156, 159, 163, 168, 172,
 188, 222, 223

G

Géographe.....27, 34
 Géométrie
 euclidienne 6, 9, 10, 68, 69, 93, 98, 132, 134, 156, 189,
 192, 194, 222
 hyperbolique.....68, 69, 70, 71, 72, 87, 130
 plan..... 71, 130, 131
 Gestion de connaissances.. 5, 6, 14, 52, 93, 156, 162, 199,
 205, 220, 223
 Grokker.....75

I

Information
 cartographique29, 30, 34, 39, 40
 graphique.....40
 Intelligence
 artificiel 105, 110, 111, 112, 160, 165
 économique 2, 18, 222
 Interaction..... 5, 8, 53, 87, 88, 89, 90, 103, 110, 120, 123,
 134, 171, 222, 225

L

Langage
 cartographie (de) 2, 11, 13, 137, 138, 163, 171
 description (de).....11, 13, 114, 138, 166, 173, 193, 194
 Lentilles
 documentaires..... 82, 83
 magiques.....89
 Logique..... 35, 105, 112, 113, 115, 160
 formelle 112, 113

M

Manipulation directe.....88
 Métaphore77
 virtuelle.....77
 Modèle de référence..... 47, 48, 49, 100
 Modélisation 13, 107, 172, 173, 179, 188, 239
 graphique..... 173, 176, 177, 182
 semi-formelle 127, 174, 176, 188
 Murs fuyants..... 82, 83, 190

O

Œil33, 38, 39, 53, 85, 243
 Ontologie
 domaine (de).....6, 9, 93, 125, 126, 135, 156, 219, 222
 Opération
 représentation (de)..... 120
 visualisation (de)..... 7, 53, 102, 189
 Ordre de cartographie..... 144, 166, 249, 251
 Organisation 2, 3, 4, 6, 7, 9, 14, 18, 19, 21, 30, 94, 95, 96,
 98, 101, 113, 127, 132, 142, 160, 161, 162, 168, 170,
 177, 182, 189, 199, 200, 201, 203, 205, 206, 207, 211,
 239

P

Paradigme
 représentation (de)..... 57, 62
 3D66, 72, 74, 77, 78, 163, 169, 225
 visualisation (de).....85, 132, 189, 194
 Perception visuelle.....30, 34, 38, 239, 243
 Point de vue théorique..... 109, 119, 179, 181
 Processus
 cartographie sémantique (de) ...2, 5, 6, 11, 93, 99, 101,
 102, 103, 104, 106, 137, 148, 163, 171, 178

R

Raisonnement distribué 32, 33, 239
 Réalité
 augmentée..... 15, 225
 Représentation
 arborescente..... 67
 avec déformation..... 10, 132, 189, 194
 en liste indentée..... 67
 externe 31, 32
 graphique6, 25, 34, 41, 44, 45, 50, 93, 94, 115, 173,
 179, 188, 239
 hiérarchie (de)73, 77, 78
 interne31, 32
 matricielle..... 66
 mental..... 6, 43, 94, 95
 nœud-lien 62, 68
 orientée relations..... 62
 orientée valeurs 56
 par pavage..... 72, 75, 76, 214
 Requêtes dynamiques88, 89
 Réseau sémantique 114, 117, 118, 122, 123, 178, 179,
 180, 182, 183, 184, 185, 186, 188

S

Sémiologie graphique34, 41
 Service de cartographie sémantique 13, 150, 152, 164,
 166, 167, 171
 Smalltalk-80 149, 151
 Société de valorisation 12, 159, 195, 223
 Structure visuelle 34, 36, 53, 57, 60, 62, 81, 88, 96, 120,
 143, 180
 Sunburst.....76
 Support externe 3, 20, 31, 32, 33, 39, 40, 52, 222
 Surcharge cognitive 73, 80, 214

T

Tables de données48, 56, 57, 59, 60, 100
 Traitement graphique 34
 Tree-Maps 72, 73, 74, 75, 128

V

Variable visuelle.....57, 59, 68
 Verrou technologique 14, 223
 Vision 208
 vision globale 4, 9, 10, 21, 46, 69, 80, 81, 94, 95, 127, 131,
 132, 189, 200, 211, 219
 Visualisation
 connaissances (de).... 4, 42, 49, 50, 51, 52, 56, 90, 127,
 240
 poly focale84, 85
 scientifique 44
 uniformes 79

8.3 La perception visuelle

L'œil est l'organe sensoriel qui nous permet de voir. Le cerveau est l'organe qui nous permet d'interpréter ces perceptions. Le passage de l'œil au cerveau correspond au processus de « perception visuelle ».

Jean-Didier Bagot définit la perception comme « l'ensemble des procédures qui nous permettent de prendre connaissance du monde environnant et de construire nos propres représentations mentales de ce monde » [Bagot, 1996].

Ainsi grâce à son sens de la vision, l'homme peut percevoir des couleurs, des formes, des ombres, des textures, etc.

Depuis quelques siècles, nous savons que la perception est rendue par l'intermédiaire d'une projection de lumière inversée au fond de l'œil. La question qui reste néanmoins en suspens est de savoir comment le cerveau comprend ces images rétiniennes. Comment le cerveau opère-t-il pour passer d'un signal lumineux à la conscience d'un objet solide, situé dans un monde extérieur, dans l'espace et dans le temps ? Depuis l'apparition de cette problématique, de nombreuses théories ont vu le jour.

D'après [Gregory, 2000], il y a 2500 ans, les philosophes grecs pensaient que nos yeux émettaient des salves de lumières pour toucher, comme des doigts, les objets. Une autre explication qui date de la même époque attribuait aux objets du monde une enveloppe, ni matérielle ni spirituelle, émanant autour des objets, comme les ronds formés par la chute d'une pierre sur une surface aquatique, conservant la forme de l'objet jusqu'à de très grandes distances. Ces « vaguelettes », nommées « données des sens », étaient supposées constituer l'intermédiaire entre les objets et la perception. Ces théories n'ont pas eu longtemps d'adeptes.

Peu de temps après, l'homme a découvert que la lumière projetait une image inversée au fond de l'œil, sur la rétine, produisant ainsi des « images rétiniennes ». René Descartes en 1637 propose, dans le cinquième discours nommé « des images qui se forment sur le fond de l'œil » extrait du livre « la dioptrique »¹ [Descartes, 1637], une expérimentation pour mettre en évidence le phénomène de projection. Ainsi il devient clair que la lumière n'entre pas, ni ne sort du cerveau. Tout ce que reçoit le cerveau comme signaux venant des organes sensoriels sont des impulsions électrochimiques ponctuelles de différentes fréquences [Gregory, 2000].

Cependant, même si l'on peut comparer l'œil à un appareil photographique, on ne peut pas résumer la perception à une production d'images rétiniennes et d'impulsions électrochimiques. En effet, ce qui frappe dans l'étude de la perception c'est la contribution importante du cerveau dans la vision. C'est lui qui crée une forte valeur ajoutée aux images rétiniennes. D'après Richard L. [Gregory, 2000] « dans des conditions idéales, la perception des objets est beaucoup plus riche que ne peut l'être une image rétinienne. La valeur ajoutée doit provenir des processus cérébraux dynamiques utilisant les connaissances stockées dans le passé afin de voir le présent et de prédire le futur immédiat ».

Il existe de nombreux modèles de ce processus qui va de la lumière à la perception des objets mais malheureusement encore à ce jour, aucun modèle ne fait l'unanimité.

¹ D'après le Trésor de Langue Française, la dioptrique est la « partie de la physique qui étudie les phénomènes de réfraction des rayons lumineux ».

Pour que l'homme puisse percevoir les données, il est nécessaire d'avoir recours à des artefacts. Dans le domaine de la visualisation d'informations, les artefacts sont nommés des « structures visuelles ». Une représentation est donc obtenue par l'action de « représenter » qui consiste à créer un ensemble de structures visuelles pour représenter les éléments de l'espace informationnel.

8.4 MDL

8.4.1 DTD de MDL

```

<!--
# Map Description Language - MDL
#
#       Current version: 2.0
#       Date: 10/08/2006
#       Author:  Christophe Tricot
-->

  <!-- Une carte est décrite par sa cartographie elle même décrite pas les connaissances à
  cartographier et les sous-cartographies-->
<!ELEMENT map (knowledge, cartographies)>
<!ATTLIST map
  name CDATA #REQUIRED
  title CDATA #IMPLIED
  version CDATA #IMPLIED
  author CDATA #REQUIRED
  description CDATA #IMPLIED
  date CDATA #IMPLIED
  persistent (true|false) #IMPLIED
>

  <!-- Les connaissances de la carte sont des schemas et des relations entre ces schemas -->
<!ELEMENT knowledge (schemas, relations, menu?)>
<!ELEMENT schemas (schema*)>
<!ELEMENT relations (relation*)>

  <!--Un schema est composé facultativement d'un menu et de description de représentations -->
<!ELEMENT schema (representationDescriptionSchema*, menu?)>
<!ATTLIST schema
  id CDATA #IMPLIED
  label CDATA #IMPLIED
  description CDATA #IMPLIED
  link CDATA #IMPLIED
  target CDATA #IMPLIED
>

  <!-- representationDescriptionSchema permet de définir une représentation spécifique identifiée
  par un numéro -->
<!ELEMENT representationDescriptionSchema (shape?)>
<!ATTLIST representationDescriptionSchema
  id CDATA #REQUIRED
  x CDATA #IMPLIED
  y CDATA #IMPLIED
  z CDATA #IMPLIED
  strokeSize CDATA #IMPLIED
  strokeColor CDATA #IMPLIED
  backgroundColor CDATA #IMPLIED

```

```

style CDATA #IMPLIED

picture CDATA #IMPLIED

font CDATA #IMPLIED
fontSize CDATA #IMPLIED
fontColor CDATA #IMPLIED
visible (true|false) #IMPLIED
permuteVisibilityOfChildrenWhenSelected (true|false|recursif) #IMPLIED
selectable (true|false) #IMPLIED
mobile (true|false) #IMPLIED
xLabel CDATA #IMPLIED
yLabel CDATA #IMPLIED
zLabel CDATA #IMPLIED
>
<!ELEMENT shape (CDATA|circle|rectangle|ellipse)?>

  <!-- Définition des différentes formes géométriques -->

  <!ELEMENT circle EMPTY>
  <!ATTLIST circle radius CDATA #REQUIRED>

  <!ELEMENT rectangle EMPTY>
  <!ATTLIST rectangle
  height CDATA #REQUIRED
  width CDATA #REQUIRED>

  <!ELEMENT ellipse EMPTY>
  <!ATTLIST ellipse
  height CDATA #REQUIRED
  width CDATA #REQUIRED>

  <!-- Une relation possède un ensemble de tuples et un ensemble facultatif de représentations -->
  <!-- De même, une relation peut posséder un menu qui sera appliqué à chacun des tuples -->

  <!ELEMENT relation (tuple*,menu?,representationDescriptionRelation*)>
  <!ATTLIST relation
    label CDATA #IMPLIED
    description CDATA #IMPLIED
    link CDATA #IMPLIED
    target CDATA #IMPLIED
    length CDATA #IMPLIED
  >

  <!-- representationDescriptionRelation permet de définir une représentation spécifique identifiée
  par un numéro -->

  <!ELEMENT representationDescriptionRelation EMPTY>
  <!ATTLIST representationDescriptionRelation
    idrepresentation CDATA #REQUIRED
    flexible CDATA #IMPLIED
    color CDATA #IMPLIED
    stroke CDATA #IMPLIED
    style CDATA #IMPLIED
    font CDATA #IMPLIED
    fontColor CDATA #IMPLIED
    fontSize CDATA #IMPLIED

```

```

visible (true|false) #IMPLIED
labelVisible (true|false) #IMPLIED
selectable (true|false) #IMPLIED

xLabel CDATA #IMPLIED

yLabel CDATA #IMPLIED
zLabel CDATA #IMPLIED
>
<!-- Un tuple est un ensemble (au moins deux) d'identifiant de schema -->
<!ELEMENT tuple (idSchema*)>
<!ATTLIST tuple
  length CDATA #IMPLIED
>
<!ELEMENT idSchema (#PCDATA)>

<!ELEMENT menu (item*)>

<!ELEMENT item EMPTY >
<!ATTLIST item
  label CDATA #REQUIRED
  action CDATA #REQUIRED
  target CDATA #IMPLIED
>

<!-- Une carte est constituée d'une représentation ou d'un ensemble de représentations répartis
suivant un modèle de disposition (layout) -->
<!ELEMENT cartographies (cartography| flowLayout | borderLayout | gridLayout |
popupLayout)>

<!-- Une représentation est obtenue en appliquant un ensemble d'opérations -->
<!ELEMENT cartography
(representationOperations,representationTransformationOperations?,visualisationOperati
ons,visualisationTransformationOperations?) >
<!ATTLIST cartography
  height CDATA #IMPLIED
  width CDATA #IMPLIED
  backgroundColor CDATA #IMPLIED
  idFocusSchema CDATA #IMPLIED
>

  <!ELEMENT representationOperations (represent | representTree)>
    <!ELEMENT represent EMPTY>
      <!ATTLIST represent idSchemaRepresentationDescription CDATA
#IMPLIED>
    <!ELEMENT representTree EMPTY>
      <!ATTLIST representTree
        idSchema CDATA #REQUIRED
        idSchemaRepresentationDescription CDATA
#IMPLIED>

  <!ELEMENT representationTransformationOperations (mapSchemaInStar |
mapSchemaInTree | mapSchemaInConeTree | forcePlacement)*>
    <!ELEMENT mapSchemaInStar EMPTY>
      <!ATTLIST mapSchemaInStar

```

```

        idSchema CDATA #REQUIRED
        length CDATA #REQUIRED
        depth CDATA #IMPLIED
    >
<!ELEMENT mapSchemaInTree EMPTY>

<!ATTLIST mapSchemaInTree

        idSchema CDATA #REQUIRED
        vSpace CDATA #REQUIRED
        hSpace CDATA #REQUIRED
    >
<!ELEMENT mapSchemaInConeTree EMPTY>
    <!ATTLIST mapSchemaInConeTree idSchema CDATA
#REQUIRED>
    <!ELEMENT forcePlacement EMPTY>

    <!ELEMENT visualisationOperations (visualisationInTouchGraph | visualisationInSVG |
visualisationInTreeBolic | visualisationInX3D | visualisationInOsEyeTree |
visualisationInOsTreeView )>
        <!ELEMENT visualisationInTouchGraph EMPTY>
        <!ELEMENT visualisationInSVG EMPTY>
        <!ELEMENT visualisationInTreeBolic EMPTY>
        <!ELEMENT visualisationInX3D EMPTY>
        <!ELEMENT visualisationInOsEyeTree EMPTY>
        <!ELEMENT visualisationInOsTreeView EMPTY>
        <!ATTLIST visualisationInOsTreeView orderBy (id|label) #IMPLIED>

    <!ELEMENT visualisationTransformationOperations EMPTY>

    <!-- Gestion des modele de disposition (layout) -->
    <!-- De gauche a droite et de haut en bas -->
<!ELEMENT flowLayout (cartography+)>

<!-- Répartition en 5 cinq zone géographique: nord sud est ouest centre -->
<!ELEMENT borderLayout (center?, north?, south?, east?, west?)>
    <!ELEMENT center (cartography)>
    <!ELEMENT north (cartography)>
    <!ELEMENT south (cartography)>
    <!ELEMENT east (cartography)>
    <!ELEMENT west (cartography)>

<!-- Répartition en tableau (ligne et colonnes) -->
<!ELEMENT gridLayout (row+)>
    <!ELEMENT row (column+)>
    <!ELEMENT column (cartography) >

<!-- Répartition en popup -->
<!ELEMENT popupLayout (popup*) >
    <!ELEMENT popup (cartography) >

```

8.4.2 Exemples de cartes avec MDL

Voici deux exemples de carte décrites en MDL et cartographiées par Os Map, notre service web de cartographie sémantique. Os Map est décrit dans la partie sur nos réalisations.

a. Exemple 1 : ontologie des êtres vivants

i. L'ordre de cartographie

```
<map name="exemple" author="totoro">
  <knowledge>
    <schemas>
      <schema id="1" label="Etres Vivants">
        <representationDescriptionSchema id="1" x="0" y="0"
backgroundColor="00FF00"/>
      </schema>
      <schema id="2" label="Vertebres">
        <representationDescriptionSchema id="1" backgroundColor="FF952B"/>
      </schema>
      <schema id="3" label="Invertebres">
        <representationDescriptionSchema id="1" backgroundColor="00B3B3"/>
      </schema>
      <schema id="4" label="Mammiferes">
        <representationDescriptionSchema id="1" backgroundColor="FFAE5E"/>
      </schema>
      <schema id="5" label="Oiseaux">
        <representationDescriptionSchema id="1" backgroundColor="FFAE5E"/>
      </schema>
      <schema id="6" label="Reptiles">
        <representationDescriptionSchema id="1" backgroundColor="FFAE5E"/>
      </schema>
      <schema id="7" label="Amphibiens">
        <representationDescriptionSchema id="1" backgroundColor="FFAE5E"/>
      </schema>
      <schema id="8" label="Poissons">
        <representationDescriptionSchema id="1" backgroundColor="FFAE5E"/>
      </schema>
      <schema id="9" label="Spongiaires">
        <representationDescriptionSchema id="1" backgroundColor="00D5D5"/>
      </schema>
      <schema id="10" label="Arthropodes">
        <representationDescriptionSchema id="1" backgroundColor="00D5D5"/>
      </schema>
      <schema id="11" label="Echinodermes">
        <representationDescriptionSchema id="1" backgroundColor="00D5D5"/>
      </schema>
      <schema id="12" label="Mollusques">
        <representationDescriptionSchema id="1" backgroundColor="00D5D5"/>
      </schema>
      <schema id="13" label="Vers">
        <representationDescriptionSchema id="1" backgroundColor="00D5D5"/>
      </schema>
    </schemas>
    <relations>
      <relation>
        <tuple>
          <idSchema>1</idSchema>
          <idSchema>2</idSchema>
        </tuple>
      </relation>
    </relations>
  </knowledge>
</map>
```

```

    <tuple>
      <idSchema>1</idSchema>
      <idSchema>3</idSchema>
    </tuple>
    <tuple>
      <idSchema>2</idSchema>

      <idSchema>4</idSchema>
    </tuple>
    <tuple>
      <idSchema>2</idSchema>
      <idSchema>5</idSchema>
    </tuple>
    <tuple>
      <idSchema>2</idSchema>
      <idSchema>6</idSchema>
    </tuple>
    <tuple>
      <idSchema>2</idSchema>
      <idSchema>7</idSchema>
    </tuple>
    <tuple>
      <idSchema>2</idSchema>
      <idSchema>8</idSchema>
    </tuple>
    <tuple>
      <idSchema>3</idSchema>
      <idSchema>9</idSchema>
    </tuple>
    <tuple>
      <idSchema>3</idSchema>
      <idSchema>10</idSchema>
    </tuple>
    <tuple>
      <idSchema>3</idSchema>
      <idSchema>11</idSchema>
    </tuple>
    <tuple>
      <idSchema>3</idSchema>
      <idSchema>12</idSchema>
    </tuple>
    <tuple>
      <idSchema>3</idSchema>
      <idSchema>13</idSchema>
    </tuple>
  </relation>
</relations>
</knowledge>
<cartographies>
  <cartography height="1000" width="1000">
    <representationOperations>
      <represent idSchemaRepresentationDescription="1"/>
    </representationOperations>
    <representationTransformationOperations>
      <mapSchemaInStar idSchema="1" length="100"/>
    </representationTransformationOperations>
    <visualisationOperations>
      <visualisationInTreeBolic/>

```

```

</visualisationOperations>
</cartography>
</cartographies>
</map>

```

ii. La carte

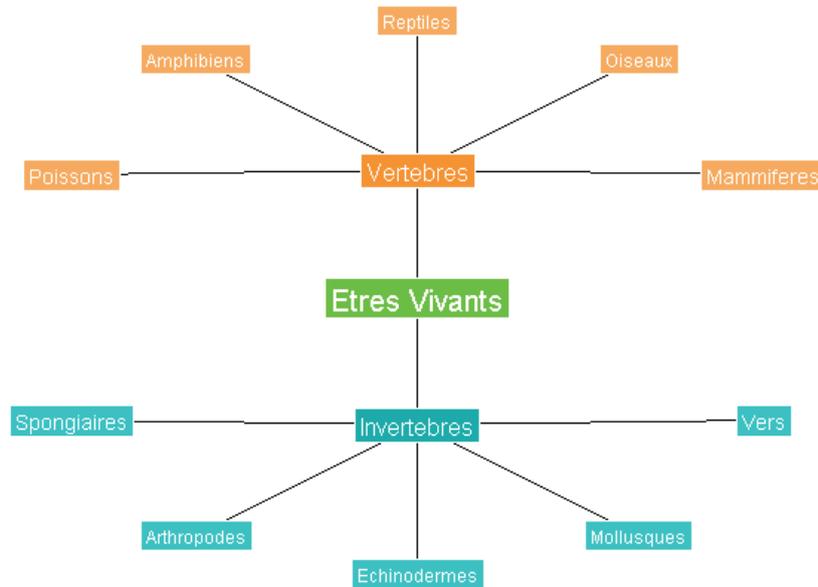


Figure 133 - MDL : exemple 1.

b. Exemple 2 : la modalité

i. L'ordre de cartographie

```

<map name="DemoModalite" author="Totoro">
  <knowledge>
    <schemas>
      <schema id="1" label="1">
        <representationDescriptionSchema id="1" x="200" y="100"
backgroundColor="777700"/>
      </schema>
      <schema id="2" label="2"/>
      <schema id="3" label="3"/>
      <schema id="4" label="4"/>
      <schema id="5" label="5"/>
      <schema id="6" label="6"/>
      <schema id="7" label="7"/>
      <schema id="8" label="8"/>
      <schema id="9" label="9"/>
      <schema id="10" label="10"/>
    </schemas>
    <relations>
      <relation>
        <tuple>
          <idSchema>1</idSchema>
          <idSchema>2</idSchema>
        </tuple>
      </relation>
    </relations>
  </knowledge>
</map>

```

```

    <tuple>
      <idSchema>1</idSchema>
      <idSchema>3</idSchema>
    </tuple>
    <tuple>
      <idSchema>1</idSchema>
      <idSchema>4</idSchema>
    </tuple>

    <tuple>
      <idSchema>1</idSchema>
      <idSchema>5</idSchema>
    </tuple>
    <tuple>
      <idSchema>1</idSchema>
      <idSchema>6</idSchema>
    </tuple>
    <tuple>
      <idSchema>1</idSchema>
      <idSchema>7</idSchema>
    </tuple>
    <tuple>
      <idSchema>1</idSchema>
      <idSchema>8</idSchema>
    </tuple>
    <tuple>
      <idSchema>1</idSchema>
      <idSchema>9</idSchema>
    </tuple>
  </relation>
</relations>
</knowledge>
<cartographies>
  <borderLayout>
    <center>
      <cartography height="300" width="400">
        <representationOperations>
          <represent/>
        </representationOperations>
        <visualisationOperations>
          <visualisationInOsEyeTree/>
        </visualisationOperations>
      </cartography>
    </center>
    <north>
      <cartography height="100" width="100">
        <representationOperations>
          <represent/>
        </representationOperations>
        <visualisationOperations>
          <visualisationInTreeBolic/>
        </visualisationOperations>
      </cartography>
    </north>
    <south>
      <cartography height="100" width="100">
        <representationOperations>
          <represent/>

```

```
        </representationOperations>
        <visualisationOperations>
          <visualisationInTreeBolic/>
        </visualisationOperations>
      </cartography>
    </south>
    <east>
      <cartography height="300" width="100">
        <representationOperations>

          <represent/>

        </representationOperations>
        <visualisationOperations>
          <visualisationInOsTreeView/>
        </visualisationOperations>
      </cartography>
    </east>
    <west>
      <cartography height="300" width="100">
        <representationOperations>
          <represent/>
        </representationOperations>
        <visualisationOperations>
          <visualisationInOsTreeView orderBy="id"/>
        </visualisationOperations>
      </cartography>
    </west>
  </borderLayout>
</cartographies>
</map>
```

ii. La carte

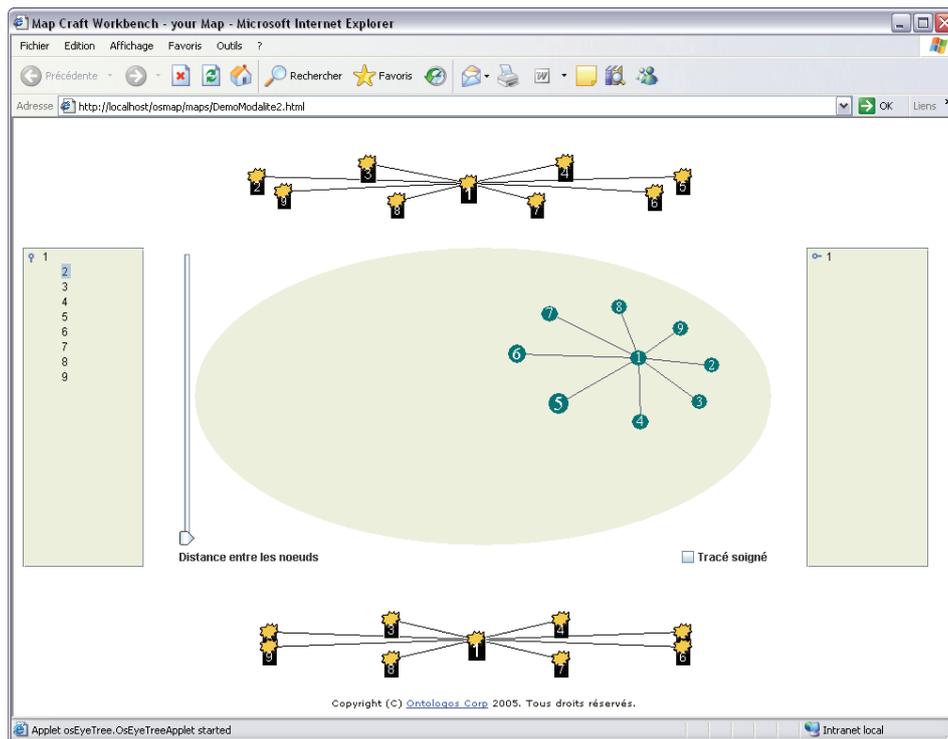


Figure 134 - MDL : exemple 2.

8.5 Description WSDL du service Os Map

```

<definitions xmlns="http://schemas.xmlsoap.org/wsdl/"
  xmlns:soap="http://schemas.xmlsoap.org/wsdl/soap/"
  xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:tns="urn:OsMap"
  targetNamespace="urn:OsMap"
  name="OsMap">
  <message name="input"/>
  <message name="output"/>

  <message name="string_input">
    <part name="in" type="xsd:string"/>
  </message>
  <message name="string_output">
    <part name="out" type="xsd:string"/>
  </message>
  <portType name="OsMapPort">
    <operation name="ping">
      <input message="tns:input"/>
      <output message="tns:string_output"/>
    </operation>
    <operation name="log">
      <input message="tns:input"/>
      <output message="tns:string_output"/>
    </operation>
    <operation name="echo">
      <input message="tns:string_input"/>
      <output message="tns:string_output"/>
    </operation>
    <operation name="map">
      <input message="tns:string_input"/>
      <output message="tns:string_output"/>
    </operation>
  </portType>
  <binding name="OsMapBinding" type="tns:OsMapPort">
    <soap:binding style="rpc" transport="http://schemas.xmlsoap.org/soap/http"/>
    <operation name="ping" selector="ping">
      <soap:operation soapAction="urn:OsMap"/>
      <input>
        <soap:body use="encoded"
  encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
        namespace="urn:OsMap"/>
      </input>
      <output>
        <soap:body use="encoded"
  encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
        namespace="urn:OsMap"/>
      </output>
    </operation>
    <operation name="log" selector="log">
      <soap:operation soapAction="urn:OsMap"/>
      <input>
        <soap:body use="encoded"
  encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
        namespace="urn:OsMap"/>

```

```
        </input>
    </output>
    <soap:body use="encoded"
encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
        namespace="urn:OsMap"/>
    </output>
</operation>
<operation name="echo" selector="echo:">
    <soap:operation soapAction="urn:OsMap"/>
    <input>
        <soap:body use="encoded"
encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
                namespace="urn:OsMap"/>
    </input>
    <output>
        <soap:body use="encoded"
encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
                namespace="urn:OsMap"/>
    </output>
</operation>
<operation name="map" selector="map:">
    <soap:operation soapAction="urn:OsMap"/>
    <input>
        <soap:body use="encoded"
encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
                namespace="urn:OsMap"/>
    </input>
    <output>
        <soap:body use="encoded"
encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
                namespace="urn:OsMap"/>
    </output>
</operation>
</binding>
<service name="OsMapService">
    <port name="OsMapPort" binding="tns:OsMapBinding">
        <soap:address location="http://localhost:8080/OsMap"/>
    </port>
</service>
</definitions>
```

8.6 Tables

8.6.1 Table des figures

Figure 1 - Exemple de carte avec l'EyeTree.....	10
Figure 2 - Exemple de carte avec le RadialTree.....	11
Figure 3 - Carte babylonienne du monde, composée probablement entre 700 et 500 ans avant Jésus-Christ (British Museum, Londres, Salle 55, Later Mesopotamia, case 15, no. 27).	26
Figure 4 - « Le géographe » de Johannes Vermeer.....	27
Figure 5 - Carte de Cassini (source IGN).	27
Figure 6 - « Carte figurative des pertes successives en hommes de l'Armée française dans la campagne de Russie de 1812-1813 » de Charles Joseph Minard (1869).	28
Figure 7 - Poser une opération : un exemple de cognition externe.....	31
Figure 8 - La cognition externe : phases du raisonnement distribué.....	32
Figure 9 - La graphique par rapport aux autres systèmes d'expression graphique.	35
Figure 10 - Familles de structures visuelles.	36
Figure 11 - Variables visuelles définies par Jacques Bertin.	37
Figure 12 - Principales interactions possibles entre les structures visuelles.	38
Figure 13 - Loi de profusion : l'œil recherche toujours le motif le plus simple.....	39
Figure 14 - Perception ambiguë.	39
Figure 15 - Cartographie de données abstraites : une activité pluridisciplinaire.....	41
Figure 16 - Exemple de cartographie : construction de métriques.	46
Figure 17 - Exemple de cartographie en visualisation d'informations [Börner et al., 2005] : partie a & b.....	46
Figure 18 - Exemple de cartographie en visualisation d'informations [Börner et al., 2005] : partie c.	47
Figure 19 - Modèle de référence de la visualisation d'informations.....	48
Figure 20 - Carte conceptuelle de l'exploration humaine [Novak, 2001].	50
Figure 21 - Croquis de Léonard de Vinci : représentation de l'idée principale du concept de machine volante.	51
Figure 22 - Les niveaux de la cartographie de données abstraites : de l'information à sa visualisation.....	54
Figure 23 - Les tables de données selon [Card et al., 1999b].....	57
Figure 24 - Seesoft : représentation d'espaces avec une structure à une dimension [Eick et al., 1992].....	58
Figure 25 - SpotFire : représentation d'espaces avec une structure à deux dimensions en nuages de points [Ahlberg, 1996].....	58
Figure 26 - Themescapes : représentation d'espaces avec une structure à trois dimensions [Wise et al., 1995].	59
Figure 27 - Table Lens : représentation « Focus+Context » d'une table de données [Rao & Card, 1994].	60
Figure 28 - Cordonnées parallèles : représentation d'espaces avec une structure multidimensionnelle [Inselberg & Dimsdale, 1990].....	61
Figure 29 - Les diagrammes en étoile : représentation d'espaces avec une structure multidimensionnelle [Chambers et al., 1983].	61

Figure 30 - Les figures de Chernoff : représentation d'espace avec une structure multidimensionnelle [Chernoff, 1973].	62
Figure 31 - Exemples de tracés de graphes de type « nœud-lien » extraits de [Nishizeki & Rahman, 2004].	63
Figure 32 - Représentation d'un graphe des cinquante mots les plus fréquents parmi les 10 % des publications les plus citées des PNAS (« Proceedings of the National Academy of Science ») [Mane & Borner, 2004].	64
Figure 33 - Représentation de réseaux sociaux [Heer & Boyd, 2005].	65
Figure 34 - représentation d'un réseau de pages Internet [Munzner & Burchard, 1995].	65
Figure 35 - "Harmony Information Landscape" [Andrews, 1995] : représentation de graphes en 3D.	66
Figure 36 - Les matrices : représentation de graphes [Ghoniem <i>et al.</i> , 2004].	66
Figure 37 - Liste indentée et le système WebTOC [Nation, 1998] : représentation d'une hiérarchie.	67
Figure 38 - Quelques algorithmes de dessin d'arbres [Hascoët & Beaudouin-Lafon, 2001].	69
Figure 39 - Les arbres de cônes : représentation d'une hiérarchie dans un espace à trois dimensions [Robertson <i>et al.</i> , 1991].	69
Figure 40 - « Circle Limit IV » de Maurits Cornelius Escher.	70
Figure 41 - Navigateur hiérarchique hyperbolique : représentation d'une hiérarchie à l'aide d'une géométrie hyperbolique [Lamping <i>et al.</i> , 1995].	71
Figure 42 - Navigateur hiérarchique hyperbolique 3D : représentation d'une hiérarchie à l'aide d'une géométrie hyperbolique [Hughes <i>et al.</i> , 2004].	72
Figure 43 - Comparaison en représentation nœud-lien et Tree-Maps.	73
Figure 44 - Tree-Maps : représentation de hiérarchies avec une technique par pavage [Johnson & Shneiderman, 1991].	73
Figure 45 - Cushion Treemaps : évolution des Tree-Maps avec une impression de relief [van Wijk & van de Wetering, 1999].	74
Figure 46 - Beamtrees : évolution des Tree-Maps en 3D [van Ham & van Wijk, 2003].	74
Figure 47 - Grokker : technique de représentation par pavage de cercles.	75
Figure 48 - Les disques d'informations : représentation par pavage d'une hiérarchie [Andrews & Heidegger, 1998].	76
Figure 49 - Sunburst : représentation par pavage d'une hiérarchie [Stasko & Zhang, 2000].	76
Figure 50 - Les cubes d'informations : représentation de hiérarchie en 3D par imbrication de cubes [Rekimoto & Green, 1993].	77
Figure 51 - Visualisation "botanique" : représentation de hiérarchies avec une métaphore virtuelle en 3D [Kleiberg <i>et al.</i> , 2001].	78
Figure 52 - Exemple de contrôle de point de vue avec deux transformations affines.	79
Figure 53 - Exemple de vues multiples : principe "overview + detail" [Jerding & Stasko, 1995].	80
Figure 54 - Astérix et Obélix : exemple de visualisation non uniforme.	81
Figure 55 - Vues non uniformes : la fonction de transformation permet de déformer les structures visuelles [Leung & Apperley, 1994].	82
Figure 56 - Vues non uniformes : la fonction et le facteur de magnification [Leung & Apperley, 1994].	82
Figure 57 - Application d'une transformation bifocale sur une et deux dimensions de la carte [Leung & Apperley, 1994].	83
Figure 58 - Lentille documentaire [Robertson & Mackinlay, 1993].	83

Figure 59 - Mur fuyant : évolution de l'affichage bifocal [Mackinlay et al., 1991].....	84
Figure 60 - Application d'une transformation poly focale [Leung & Apperley, 1994].....	84
Figure 61 - Tableaux optiques : application de plusieurs points de focalisation sur un tableur [Rao & Card, 1994].	85
Figure 62 - Fisheye filtrant : application du paradigme du fisheye (dessin de droite) sur un arbre (dessin de gauche) [Hascoët & Beaudouin-Lafon, 2001].	86
Figure 63 - Fisheye déformant : application du paradigme du fisheye (dessin de droite) sur les coordonnées cartésiennes des éléments d'un graphes (dessin de gauche) [Sarkar & Brown, 1992].	86
Figure 64 - les interactions dans la cartographie.....	87
Figure 65 - Exemples de répartitions possibles des représentations des collaborateurs de l'organigramme.	95
Figure 66 - Représentation de l'organigramme avec une technique de type Treemaps. .	96
Figure 67 - Les niveaux de la cartographie de données abstraites : de l'information à sa visualisation.....	100
Figure 68 - Notre processus de cartographie sémantique.....	104
Figure 69 - Illustration de notre processus avec un organigramme.	105
Figure 70 - Changements de nature d'espace possibles lors d'une opération.....	106
Figure 71 - SNDF pour décrire des modélisations et des espaces informationnels <i>structurés</i>	117
Figure 72 - SNDF : notre formalisme permet le partage d'objets entre deux modélisations.	120
Figure 73 - Les cartes ont des structures visuelles similaires.	121
Figure 74 - Les éléments possèdent une liste de descriptions de représentations.	122
Figure 75 - Positionnement de la problématique par rapport à notre processus.	125
Figure 76 - Indexation des documents sur une ontologie de domaine.	126
Figure 77 - Prototype 1 : représentation en liste indentée (treeview simple).	129
Figure 78 - Prototype 2 : les arbres de cônes.	130
Figure 79 - Prototype 3 : les arbres hyperboliques.....	131
Figure 80 - EyeTree : opérations sur l'espace informationnel <i>structuré</i>	132
Figure 81 - EyeTree : interactions avec la carte.....	133
Figure 82 - EyeTree : exploration d'une structure de concepts.	133
Figure 83 - RadialTree : exploration d'une structure de concepts.	135
Figure 84 - Exemple d'une carte contenant un organigramme.	137
Figure 85 - Structure d'une cartographie décrite en MDL.	139
Figure 86 - Gestionnaires de mise en page de MDL.	142
Figure 87 - Exemple de carte avec MDL.	145
Figure 88 - Architecture courante pour la cartographie sémantique.....	148
Figure 89 - Principe de notre proposition d'architecture de cartographie sémantique.	148
Figure 90 - Principe de MVC.	149
Figure 91 - MVC : distribution du modèle et de ses vues.	150
Figure 92 - Configuration possible de notre architecture de cartographie sémantique distribuée.	153
Figure 93 - Mécanisme des communications : illustration avec un diagramme de séquences.	154
Figure 94 - KCW et la démarche Ousia.....	161
Figure 95 - Os Map construit des cartes selon différents formats.	163
Figure 96 - Usage d'Os Map.....	164
Figure 97 - Principe des services web.	165

Figure 98 - Os Map : cartographie d'un système de fichiers.....	169
Figure 99 - Os Map : cartographie d'un référentiel de compétences.	170
Figure 100 - Os Map : cartographie d'une pyramide des âges.....	171
Figure 101 - Modélisation graphique avec <i>Protégé</i>	173
Figure 102 - OCW : cartographie d'ontologie.....	174
Figure 103 - Exemple de carte heuristique.	175
Figure 104 - Exemple de carte conceptuelle sur la biodiversité.....	175
Figure 105 - SNCW : besoins pour la modélisation graphique.....	177
Figure 106 - Principe de SNCW.	178
Figure 107 - SNCW : modélisation de satellite.....	179
Figure 108 - SNCW : plusieurs vues d'une même modélisation.	180
Figure 109 - SNCW : attributs de représentation.	181
Figure 110 - SNCW : création de représentations.	181
Figure 111 - SNCW : partage d'objets entre les modélisations.....	181
Figure 112 - Architecture de SNCW.	182
Figure 113 - Interface de SNCW.	187
Figure 114 - EyeTree : caractéristiques de la fonction d'amplification.....	190
Figure 115 - EyeTree : résultat de l'application de la fonction d'amplification.	190
Figure 116 - Exemple d'intégration de l'EyeTree dans KCW.....	191
Figure 117 - RadialTree : déformation dans la zone de focus.	192
Figure 118 - RadialTree : effet de type menu fisheye.....	193
Figure 119 - Exemple d'intégration de RadialTree dans KCW.....	194
Figure 120 - Exemple de carte annuaire.....	202
Figure 121 - Exemple de carte des emplois accessibles.	204
Figure 122 - Cartes prévisionnelles : identification des compétences critiques.....	206
Figure 123 - Carte stratégique : pyramide des âges des collaborateurs par bureau.....	207
Figure 124 - Cartes stratégiques : vision globale.....	208
Figure 125 - Entreprise & Carrières : extraits.	208
Figure 126 - Le système Lighthouse.	213
Figure 127 - Recherche de documents avec <i>OntoExplo</i>	215
Figure 128 - <i>OntoRequest</i> : construction de requêtes ontologiques.....	215
Figure 129 - <i>OntoRequest</i> : exemple de requête.	216
Figure 130 - <i>OntoRequest</i> : sélection d'un concept.	216
Figure 131 - <i>Misti</i> : recherche classique.	218
Figure 132 - <i>Misti</i> : recherche avec l'EyeTree.....	218
Figure 133 - MDL : exemple 1.	251
Figure 134 - MDL : exemple 2.	253

8.6.2 Table des tableaux

Tableau 1 - Domaines de signification des dessins.....	35
Tableau 2 - Synthèse des principales variables visuelles.	37
Tableau 3 - Paradigmes d'interaction au niveau de l'espace informationnel.....	88
Tableau 4 - Paradigmes d'interaction au niveau de l'espace informationnel <i>représenté</i> ... 88	
Tableau 5 - Paradigmes d'interaction au niveau de l'espace informationnel <i>visualisé</i>	89
Tableau 6 - Classes d'opérations en MDL.	143
Tableau 7 - Agents logiciels de notre architecture.	152

8.7 Crédits images

8.7.1 Les cités obscures

Toutes les images placées au début de chaque partie sont extraites d'une collection d'ouvrages intitulée « Les Cités Obscures », collaboration du dessinateur belge de bande dessinée François Schuiten et de son ami, le scénariste Benoît Peeters. Elles sont reproduites avec l'aimable autorisation de Benoît Peeters.

8.7.2 Circle Limite IV

L'image « Circle Limit IV » de M.C. Escher est reproduite avec l'aimable autorisation de « The M.C. Escher Company – the Netherlands ».





Résumé

Cette étude a débuté avec un constat important : les organisations sont submergées par le nombre toujours croissant d'informations qu'il est nécessaire de maîtriser pour mener à bien leurs activités. En effet, étant obligées de faire face à la mondialisation des flux et des services, les organisations ont découvert de nouveaux besoins et elles ont développé de nouvelles activités : veille stratégique, gestion de l'innovation, capitalisation des savoirs, gestion du capital intellectuel et humain, intelligence économique, etc. Depuis, elles collectent et stockent toujours plus d'informations et leur espace informationnel ne cesse de grandir. Dans ce contexte mondialisé, l'avenir des organisations passe par la maîtrise de leur espace informationnel.

La maîtrise d'un espace est une problématique universelle qui préoccupe l'homme depuis son origine. La cartographie résulte de cette problématique ; elle permet de visualiser un espace pour mieux l'appréhender. Chaque carte joue alors le rôle de support externe de la pensée pour amplifier la cognition des personnes qui l'utilisent. L'étude de la cartographie nous a appris que pour construire des cartes à partir de l'espace informationnel des organisations, il est nécessaire de mettre en œuvre des paradigmes de représentation, de visualisation et d'interaction. Cette « recette » est maintenant bien connue et maîtrisée. Cependant elle ne répond que partiellement à nos besoins. En effet, il subsiste deux problématiques à résoudre pour y parvenir : comment proposer une cartographie basée sur la sémantique du domaine ? comment déterminer les paradigmes à mettre en œuvre pour fournir des cartes adaptées à l'utilisateur ?

Face à cette double problématique, nous avons défini une nouvelle cartographie capable de construire des cartes adaptées aux besoins des organisations. Nous la nommons « cartographie sémantique ». Elle repose sur nos différentes propositions et sur leurs réalisations. La mise en œuvre de la cartographie sémantique dans le monde des entreprises nous a permis de valider l'ensemble de l'étude mais surtout de s'assurer que la problématique était bien résolue.

Mots-clés : Cartographie d'informations et de connaissances, visualisation d'information, processus de cartographie, ontologies, représentation des connaissances, intelligence artificielle, gestion et recherche d'informations graphiques, architectures à agents.

Abstract

This study has begun with an important observation: organizations are more and more submerged by information whose management is necessary for their activities. As a matter of fact, the globalization of flows and services has created new needs and then new activities have been developed: strategic intelligence, innovation management, knowledge capitalization, intellectual and human capital management, economic intelligence, etc. Insofar as more and more information are collected, the enterprise's informational space does not stop to increase. In such a context, the future of organizations goes through the control of their informational space.

The control of a space is a universal problem which concerns mankind since its origin. Cartography results from this problem. It makes it possible to visualize a space for a better apprehension. Each map plays the role of an external medium for mind in order to amplify the user's cognition. The study of cartography has revealed that building maps from the informational space of organizations requires to carry out paradigms of representation, visualization and interaction. This process is now well known and controlled. However it meets only partially our needs. Indeed, it remains two problems to solve to reach that point: how to take into account the domain's semantics and how to select the different paradigms in order to provide the most suitable maps for users?

Faced with this two-fold problem, we have defined a new cartography paradigm for building maps which fit to the needs of organizations. We have named it "semantic cartography". This new cartography paradigm relies on our different propositions and realizations. Our proposition has been validated through different industrial applications, e.g. content management systems and competency management systems.

Keywords: Cartography of information and knowledge, information visualization, cartography process, ontologies, knowledge representation, artificial intelligence, graphic management and search of information, Multiagent system.