



**Université de Technologie de Troyes**

**Thèse**

**pour obtenir le grade de**

**Docteur de l'Université de Technologie de Troyes**

**Informatique**

**Présentée et soutenue publiquement par**

**Samuel PARFOURU**

**le 13 mars 2008**

**Approche méthodologique et instrumentale de  
construction et de restitution de documents structurés**

*Une étude empirique dans un contexte de gestion des connaissances au sein de  
EDF Recherche et Développement*

*Directeurs de thèse :*

**M. Manuel Zacklad**, Université de technologie de Troyes, Professeur des universités

**M. Sylvain Mahé**, EDF Recherche et Développement, Ingénieur Chercheur

*Rapporteurs :*

**M. Jean Charlet**, *Rapporteur*, Université Paris 8, Professeur des universités

**M. Khaldoun Zreik**, *Rapporteur*, Université Paris 8, Professeur des universités

*Examineurs :*

**M. Pascal Salembier**, Université de technologie de Troyes, Professeur ECC HDR

**M. Benoit Ricard**, EDF Recherche et Développement, Ingénieur Chercheur

**M. Mauro Gaïo**, Université de Pau et des Pays de l'Adour, Professeur des universités

« Ce n'est pas parce que c'est difficile que nous n'osons pas. C'est parce que nous n'osons pas que c'est difficile. » (*Sénèque*)

*A cette petite famille du pays des « Agytés »,*

*A Karen qui pendant longtemps a été ma première lectrice*

## Remerciements

Je me souviens, c'était en décembre 2004. J'arrivais sur l'Ile de Chatou pour prendre possession de mon bureau et la plupart des gens partaient en vacances de Noël, laissant derrière eux, études et expériences. Nous étions une petite poignée pendant ces quelques jours, accolés aux fêtes de fin d'année, mais cela n'a pas duré longtemps. Le travail de thèse est, de mon point de vue, une aventure humaine pendant laquelle on côtoie, rencontre et découvre de nombreuses personnes qui, tout au long de ces trois années, ou seulement l'espace d'un moment, prennent un rôle important pour mener une recherche que l'on souhaite de qualité.

*Une thèse, c'est avant tout de l'inspiration...*

... elle représente un engagement personnel et professionnel très long. Aussi, je dois bien admettre qu'il y a un long cheminement pour en arriver là. Un regard rétrospectif me pousse à évoquer ici des gens qui m'ont fait aimer le monde de la recherche, qui m'ont transmis leur passion, en particulier lors de mon passage à l'université de Caen. Aussi, je remercie chaleureusement *Bruno Cremilleux, Nadine Lucas, Mauro Gaïo, Jacques Madeleine, Stéphane Ferari* et bien sûr *Khaldoun Zreik* rencontré(s ?) à cette époque. Ils sont encore aujourd'hui des soutiens fidèles et restent de bons conseillers malgré mon départ de cette université.

J'ai aussi une pensée pour *François Rioult*, à l'époque Doctorant, qui, dans une conversation de couloir, m'avait incité à porter un regard (attentif ? prudent ? nouveau ? regard tout court ne suffit pas : il faut lui adjoindre un qualificatif) sur le métier de chercheur : je l'en remercie ; il ne s'était pas trompé, il est aujourd'hui maître de conférence et c'est une chance pour les étudiants qui ont l'occasion de croiser son chemin.

*Une thèse, c'est aussi un cheminement qui s'accompagne d'un encadrement...*

Aussi, je tiens à remercier *Manuel Zacklad*, directeur de thèse, qui m'aura fait confiance et m'aura accompagné dans cette entreprise. Je me souviens encore de notre premier coup de fil et de notre première rencontre : de brefs moments, il n'en saurait être autrement mais qui, comme tout au long de ces trois années, ont été ponctués d'une grande sympathie et d'une efficacité certaine.

Je remercie également *Sylvain Mahé* qui a conduit l'encadrement au sein d'EDF. Je ne peux que souligner le soin qu'il a apporté pour que cette thèse se déroule dans les meilleures conditions. Pendant longtemps, nous disposions de visions contradictoires quant aux axes d'orientation à donner à l'étude. J'espère toutefois qu'au final, nous aurons trouvé un terrain commun. Je reste admiratif devant son souci du détail si ancré, si prononcé.

Je dois également évoquer ici *Benoît Ricard*, membre du comité technique de cette thèse. Il est important pour moi de mettre l'accent sur l'apport de cet homme, si discret, si modeste, à la fois doué d'un esprit de synthèse dont on ne peut que s'inspirer et d'une élégance d'écriture devant laquelle on ne peut être qu'admiratif. C'est au hasard d'une question, d'une remarque qu'il vous oriente vers le point à améliorer, à expliquer, à approfondir. Je le remercie d'avoir investi, très tôt, une partie de ma démarche dans ces propres travaux.

*Passage obligé, après trois ans, une thèse, c'est une évaluation...*

Elle passe tout d'abord entre les mains et sous le regard de deux rapporteurs. J'exprime en premier lieu, mes sincères remerciements à *Jean Charlet*, qui, bien que n'ayant que peu de connaissances sur mon travail, a accepté d'y apporter toute son attention. Je remercie aussi, une fois encore, *Khaldoun Zreik*, qui suit mon parcours et encourage mes efforts depuis maintenant cinq années, et qui me fait l'honneur de relire et de rapporter patiemment ce mémoire.

Et puis, c'est également toute l'attention des examinateurs qui sera mise à contribution. J'exprime ma reconnaissance à *Pascal Salembier* et de nouveau à *Benoît Ricard* d'avoir accepté d'être membres du jury.

*Une thèse ne saurait se passer de financement...*

C'est « le nerf de la guerre ». Aussi, je remercie d'abord la société EDF (Electricité de France) qui est à l'origine du financement et ensuite, la division Recherche et Développement pour m'avoir accueilli au sein du département STEP. Je remercie également *Philippe Klein* et *Karine Feissart* pour avoir rendu possible la prolongation de trois mois de cette thèse.

*Un thèse passe également par de la confiance...*

Aussi, j'exprime ma gratitude à *Valery Just*, chef de groupe, qui a participé à mon recrutement : malgré certains doutes quant à mes axes d'étude, à un stade où je les exprimais sûrement maladroitement, il m'a accordé une confiance certaine et m'a laissé une réelle latitude dans mes orientations.

Cette confiance a manifestement été reconduite par *Arnaud Fourniguet* qui a succédé à *Valery Just*. J'imagine la difficulté de s'imprégner d'une démarche de thèse déjà commencée depuis presque deux ans. Je salue sa confiance pour nous avoir laissé poursuivre notre étude telle que nous l'avions engagée.

Il me faut également souligner le rôle important des chefs de projets sur lesquels s'appuie notre étude. En premier lieu, *Catherine Bauby* qui, très tôt, a accordé un intérêt à mes propositions et qui a ouvert les portes de son propre projet.

Je remercie aussi *Philippe Nonclerq* pour son regard intéressé à l'égard de mon travail. Il nous aura autorisé à bouleverser radicalement la concrétisation des guides de diagnostic, me donnant l'occasion de me rendre sur le terrain pour être au contact des experts et des opérationnels.

*Durant une thèse, vous cherchez à susciter l'intérêt...*

Merci à *Nada Matta*, qui aujourd'hui a choisi une autre vocation, mais qui a exprimé de la ferveur lorsque, pour la première fois, je lui ai exposé mon projet et qui m'a gratifié de ses encouragements durant notre déplacement dans ce pays lointain qu'est l'Inde. Je garderai en mémoire ce modèle de gentillesse et de générosité.

Je dois également évoquer l'intérêt que *Sabine Mas* a porté à mon travail, durant un voyage en train de Troyes vers Paris, alors qu'elle me connaissait à peine. Elle m'a ouvert les

portes de l'Ecole de Bibliothéconomie et des Sciences de l'Information de Montréal. Mon intervention dans le cadre d'une conférence midi m'a beaucoup apporté.

*Sabine Mas* m'a donné l'occasion de rencontrer *Yves Marcoux*, chercheur éclairé. Nos discussions ont été essentielles dans ma démarche : je ne peux que le remercier pour nos échanges dans un Second Cup à Montréal, ou bien lors de la 9<sup>ème</sup> édition de la conférence CIDE à Fribourg.

Je remercie aussi les commanditaires et les experts, mais également *Bruno Charbonnier* et *Emmanuel Remy*, chez qui, le travail produit dans le cadre de cette thèse a pu susciter de l'intérêt.

*3 ans de thèse, ce sont également des rencontres...*

On baigne dans un environnement où l'on évolue et on travaille sur des problématiques plus ou moins connexes à la démarche de recherche. Pour ma part, l'environnement se situe tant à Troyes qu'à Chatou.

Aussi, je remercie les membres du laboratoire TechCICO de l'UTT pour leur sympathie. Les doctorants qui ne sont pas en permanence au laboratoire bénéficient d'un accueil et d'un soin particulier malgré le caractère très ponctuel et intermittent de leurs visites.

Je souligne l'apport crucial du travail de *Stéphanie Clément* et *Christine Danau*. Nous serions en effet bien orphelins dans les démarches administratives et dans les préparations de conférence sans leur concours.

J'adresse ma profonde gratitude à *Heddie Zaher* et *Chaker Diaz* qui ne peuvent être décrits que comme des modèles d'altruisme.

Du côté de la capitale, un petit clin d'œil aux membres des compétences en traitement du signal, système dynamique et ingénierie des connaissances du département P1B. L'ambiance chaleureuse, parfois même familiale, de ce groupe de recherche a joué un rôle important pour mener à bien ce travail.

Un merci particulier à *Fabien Peisey*, qui aujourd'hui vogue vers d'autres horizons d'étude, mais qui a apporté une forme de sérénité qu'il véhiculait dans ce bureau ponctuée de profondes conversations ou de plaisanteries.

Je remercie également *Guy Durso*, homme d'expérience, qui possède de profondes qualités humaines et dont la porte a toujours été ouverte pour m'offrir ses précieux conseils.

Et que dire de la petite *Mabrouka El-Guedri*, rencontrée alors qu'elle était stagiaire, devenue doctorante, et aujourd'hui titulaire. Son sourire rayonnant aura beaucoup égayé cette longue période. Je lui souhaite de finir au mieux sa thèse et de bénéficier d'une belle carrière chez EDF.

*Au-delà des rencontres, une thèse c'est également de l'implication...*

Une implication personnelle, mais également l'investissement de tierces personnes. Aussi, je ne sais au combien remercier *Alain Grassaud*, sociologue mécanicien, qui aura fait

preuve d'une implication très importante. Il nous aura offert un terrain réel de mise en pratique de l'approche que nous proposons, tout en jouant un rôle majeur dans son développement.

Soulignons également l'implication de *Philippe Haïk*, « co-bureau » et ami, soutien du premier jour, qui a toujours cru en mon travail, m'encourageant et s'y investissant. Il représente un maillon essentiel dans la communication de ce travail au sein de l'entreprise, me permettant de rencontrer d'autres maillons essentiels dans cette thèse. Je lui dois également beaucoup pour mon avenir professionnel. J'espère que nous pourrions poursuivre nos échanges et nos réflexions malgré des bureaux géographiquement éloignés.

*La fin de thèse peut être vue comme une issue, mais c'est plutôt un commencement...*

Pendant et en fin de rédaction, on prend conscience de tout ce qu'il reste à faire. Pour ma part, je vais avoir l'opportunité de poursuivre mon activité au sein d'EDF. Je remercie ici *Laurent Coudert*, *Evelyne Leray*, *Germain Pot* et enfin *Benoît Richard* qui ont accordé un intérêt à ma candidature et me donnent l'opportunité de poursuivre mon travail dans un environnement offrant un cadre stimulant et de nombreux terrains d'investigations vierges.

*Au cours de ces trois ans de travail, on a aussi besoin d'un peu de renfort...*

La thèse n'est pas le travail d'un seul homme ; beaucoup de gens viennent y apporter une contribution. Dans certains moments, on cherche du renfort pour approfondir l'approche et par exemple, la mettre en œuvre. Je citerai ici l'apport en juin 2007 de mon groupe d'étudiants de LO02 de Troyes qui, dans le cadre d'un projet à réaliser, aura contribué à l'approche de restitution.

Il faut également souligner le travail remarquable de *Jérôme Gagnière*, venu renforcer le développement des guides de diagnostic durant 6 mois dans le cadre d'un stage d'ingénieur. Merci pour la confiance dont il a fait preuve dans sa montée en compétence sur mon approche méthodologique ainsi que sur les technologies utilisées, mais également pour son intérêt et sa grande implication dans son stage. Il possède cette qualité de « l'amour du travail bien fait ». J'espère qu'il me pardonnera pour ma maigre disponibilité.

De l'aide, il en faut également dans l'écriture du mémoire. Plus de 300 pages, ça paraît interminable, voire même fastidieux à lire et relire et corriger. Je remercie *Benoît*, *Sylvain*, *Philippe* et *Manuel* pour y avoir porté un regard scientifique.

J'exprime ma plus profonde gratitude aux personnes ayant contribué à écarter les fautes de frappe, les fautes d'orthographe et de grammaire, l'amélioration de l'expression. Merci beaucoup à *Vincent* qui m'a accordé beaucoup de temps au détriment de ces lectures habituelles. Merci également à *Karen*, longtemps la première lectrice de mes rapports. Merci à *Marie* qui trouvant une partie du mémoire sur une table familiale aura apporté sa part de correction et enfin à *Albannie* qui m'aura offert son œil éclairé sur les dernières parties rédigées.

*Pendant 3 ans, on cherche parfois des refuges...*

La période de doctorat est parfois pesante. Elle est parsemée de moment d'euphorie mais également de période de doutes où l'on se sent plutôt au creux de la vague. Pour certains

de ces moments, je me suis réfugié à Montréal et c'est *Sabine* qui m'a généreusement accueilli, m'offrant un havre de paix, lors de ces coupures sur l'autre continent. Je pense que nous garderons bien longtemps en mémoire certaines soirées avec la Croix rouge de Montréal ou à regarder les Canadiens pour leur passage en série. Je remercie également *Christine Dufour* pour son accueil lors de mon passage au printemps 2006. Je reviendrai me confronter à certains casse-têtes de ton salon.

J'exprime ma reconnaissance à *Karmele*, pour les quelques jours passés en Espagne au printemps 2007. Les moments passés au Corrilo m'ont permis de beaucoup m'aérer l'esprit.

Je remercie également *Nicolas* ainsi que sa femme *Emilie* de m'avoir ouvert la porte de leur maison l'espace d'un week-end ou d'une journée.

*Au delà des refuges, on cherche de l'écoute*

Je remercie *Claire* qui a repris contact avec moi à son retour des Etats-Unis. Nos déjeuners ou dîners à Cabourg autour d'un plat de moules frites accompagné d'un breuvage adapté, ont été des moments importants.

Petit clin d'œil à *Nadine*, à la fois pour m'avoir ouvert les portes du séminaire I3 du laboratoire GREYC, mais également pour notre échange durant un menu « Saint Jacques » à Courseulles.

J'adresse encore ici toute ma gratitude à *Kaldhoun* pour son écoute et ses conseils, en particulier lors des vernissages à la galerie Europa.

*Et lorsque l'écoute ne suffit plus, il faut du soutien...*

Je l'ai déjà évoqué trois ans de thèse, c'est long. Comme l'on dit souvent, ce n'est pas tant de réaliser le travail qui est difficile que de tenir particulièrement moralement. Il y a de nombreux noms que je pourrais mentionner, puisque beaucoup de gens ont participé à l'édifice sans nécessairement toujours s'en rendre compte. J'évoquerai tout particulièrement *Cathy*, aujourd'hui maman d'un petit *Antoine*, à laquelle je suis infiniment reconnaissant : tu as été un soutien, une confidente, une partenaire de délires en tout genre ; je pense que tu m'as souvent détourné d'un sentiment de découragement.

*Durant cette période, ce sont également les amitiés de longue date qui comptent...*

Ce sont ces personnes présentes malgré l'éloignement. Au-delà des remerciements, ce sont plus des excuses que je dois faire ici : pendant ces trois ans, j'aurais mis un peu de distance, voir même peut-être un peu de négligence, non de manière délibérée, avec famille et amis. Certains amis se sont éloignés, mais beaucoup sont restés et je ne peux qu'avoir une pensée pour eux et souligner la chance de les avoir. Merci à *Mathilde*, *Yves*, *Marilyne* mais également *Jean-Yves*, *Odile*, *Agnès*, *David*...

*En écrivant ces lignes, je pense également à ma famille...*

Ils sont toujours aussi présents malgré mon éloignement et parfois ma prise de distance. Je suis fier d'avoir ces cinq personnes, et puis deux de plus et encore une petite

dernière *Anais* à la fin de l'année passée. Merci à *Marie, Jacques, Vincent, Karen, Régis, Séverine* et *James*.

*Il y aurait encore pleins d'autres choses à évoquer...*

Et donc, pour terminer, je ne saurais écarter l'importance dans la finalisation de ce travail de cette amoureuse des livres qui m'accompagne depuis ces derniers mois. Et à l'image de ce personnage de la société de production Aardman, à toutes les personnes ayant participé de près ou de loin à ce travail, cité ou non dans ces remerciements, je dis : « We did it !! ».

## Résumé

L'étude portée dans ce mémoire s'intéresse aux conditions de réalisation de l'inscription d'une connaissance sur un support. Nous avons ainsi mené une étude empirique dans un contexte de gestion des connaissances en milieu industriel au sein d'EDF - Recherche et Développement. Nous avons ainsi pris pour cadre d'analyse le document au travers des 3 dimensions qui le caractérise : medium, signe, forme.

Ces trois dimensions ont permis d'investiguer le contexte social de l'activité de gestion des connaissances, la construction de sens entre les acteurs ainsi que la forme des informations inscrites. Nous nous sommes en particulier, attardés sur les outils et supports exploités pour inscrire et faire transiter les informations, en mettant l'accent sur l'importance de la forme documentaire comme vecteur de traitement et de diffusion de l'information.

De cette analyse, nous avons développé une approche méthodologique et instrumentale de construction et de restitution de documents structurés. Le choix du document structuré et le développement de notre approche est motivé par les besoins liés à notre terrain d'étude. Notre approche trouve ses fondements dans l'articulation de l'ingénierie documentaire, l'interface homme machine et la sémiologie graphique. Cette dernière a ainsi été adaptée et étendue pour permettre la construction d'un document, mais également pour prendre en considération les capacités des logiciels existants en terme de restitution.

Cette approche a été confrontée à 2 cas d'application réels, nous permettant de définir les grandes lignes de développement d'un Atelier sémiotique pour assister la construction des documents structurés.

## Mots Clés

Document électronique, Document structuré, Support numérique, Sémiologie graphique, Ingénierie documentaire, Ingénierie des connaissances, Technologie de l'information, Méthode, Restitution, Structuration d'information

## Abstract

This study focus on the process of writing knowledge on physical or digital support. We lead an empiric study in the field of knowledge management in an industrial context. Our analyze is guided by the three dimensions that are currently associated to documents : medium, sign and form.

This three dimensions gave us the opportunity to investigate the social context of knowledge management, the creation of sense between actors and the way to formalize information. We identified instruments and supports (physical or digital) that are used to broadcast information. In particular, we noted the importance of of documents for the analyse and the broadcast of information.

From this analyze, we have developped a methodology and tools for the design and the restitution of structured documents. The choice of structured documents as a support for our approach has been motivated by the needs of in-(context knowledge management. Our approach is a bridge between the following research fields : document, human computer interface and graphic semiology. The graphic semiology has been adapted in order to allow the design of documents and in order to take into account the limited restitution functiuns of existing softwares.

Our approach has been applied on two real case studies. That is permeted to define the outline of a semiotic software framework to support the design of structured documents.

## Keywords

Electronic document ; Structured document ; Electronic support ; document engineering ; Graphic semiology ; Knowledge engineering ; Information technology ; Method ; Restitution ; Information structuring

# Sommaire

Introduction .....	16
Partie I - Problématique .....	20
Chapitre 1 - Le terrain.....	21
1 - Gestion des connaissances .....	21
1.1. La gestion des connaissances à la croisée de 3 dimensions.....	21
1.2. Quelques méthodes.....	24
2 - Instrumenter la Gestion des Connaissances : l'Ingénierie des Connaissances.....	25
2.1. Des systèmes experts à l'orientation dans les supports numériques .....	26
2.2. L'avènement des supports numériques.....	28
2.3. Les produits de l'ingénierie des connaissances.....	33
3 - EDF et la gestion des connaissances en milieu industriel.....	39
3.1. Un enjeu majeur .....	39
3.2. Quelques exemples.....	43
Chapitre 2 - Problématique .....	49
1 - L'objet d'étude .....	49
2 - Les problématiques.....	52
2.1. L'activité de capitalisation de connaissances .....	52
2.2. Restitution d'informations .....	55
2.3. Orientations de notre étude.....	60
Partie II - Approche théorique.....	62
Chapitre 3 - Ingénierie documentaire .....	63
1 - Du document au document numérique structuré.....	63
1.1. Le document comme preuve.....	63
1.2. Le document numérique .....	64
1.3. Le document numérique structuré.....	65
2 - La réutilisation face au concept de document .....	71
2.1. Les reformulations du numérique .....	72
2.2. La réutilisation face au concept de document.....	76
3 - Le document pour l'action .....	77
3.1. Définition.....	77
3.2. Construction des documents .....	78
4 - Du document à l'interface homme machine.....	81

Chapitre 4 - Ingénierie des Interfaces Homme Machine .....	83
1 - Interaction Homme Machine.....	83
2 - De l'interaction homme machine à la métacommunication.....	87
2.1. Dialogue homme machine .....	88
2.2. Sémiotique et IHM .....	92
2.3. L'interaction homme machine vue comme une meta communication.....	96
3 - Interface homme information .....	99
Chapitre 5 - Ingénierie de la forme.....	102
1 - Visualisation d'informations .....	102
1.1. Classification des techniques de visualisation .....	103
1.2. Visualisation et interaction .....	108
1.3. Conception d'une visualisation.....	110
2 - La sémiologie graphique (Bertin 1973) .....	116
2.1. Contenu et contenant d'une représentation graphique .....	117
2.2. Méthode de construction.....	120
2.3. Au-delà de Bertin .....	123
3 - Ingénierie de la forme, de l'interface et du document .....	124
Partie III - Analyse empirique des systèmes de gestion de connaissances .....	127
Chapitre 6 - Sous la dimension medium .....	130
1 - Regard sur l'exemple des guides de diagnostic au travers de la dimension medium... 130	
1.1. L'analyste en médiateur.....	131
1.2. Des acteurs en interaction médiatisée.....	132
1.3. Outils exploités pour la capitalisation et la diffusion .....	133
2 - Activité de gestion des connaissances au regard de la dimension médium .....	134
Chapitre 7 - Sous la dimension signe.....	137
1 - Définition du signe dans le développement des guides de diagnostic .....	138
1.1. Formalisation multi-domaines et multi-points de vue.....	138
1.2. Le document comme support de capitalisation.....	139
2 - Analyse de méthodes et outils d'aide à la gestion des actifs de production sous l'angle du signe .....	142
2.1. Formalisation computationnelle.....	142
2.2. La mise à disposition des résultats de sortie du système.....	143
3 - Activité de gestion des connaissances au regard de la dimension signe.....	144
3.1. Face à la réalisation d'un document .....	145

3.2. Face à un système à base de connaissances .....	145
Chapitre 8 - Sous la dimension forme.....	148
1 - Analyse des documents produits .....	148
1.1. Word et la production d'un contenu orienté présentation .....	148
1.2. Excel au cœur d'une pratique de structuration de l'information.....	151
1.3. Du « WYSIWYG » au « WYSIWYM » : au-delà de la capitalisation du fond, capitaliser la forme.....	152
2 - Analyse de la forme dans les outils d'aide à la gestion des actifs de production.....	154
2.1. Le formulaire.....	154
2.2. Des sources de données dissociées dans la restitution des informations.....	156
3 - Activité de gestion de connaissances au regard de la dimension forme .....	157
Chapitre 9 - Conclusion d'analyse.....	160
1 - Caractérisation de l'activité de gestion des connaissances .....	160
1.1. Un registre d'écriture particulier pour les documents de capitalisation .....	160
1.2. Une diversité d'outils mobilisés dans la création des documents.....	160
1.3. L'importance de l'autonomie de l'analyste .....	161
1.4. Rédaction d'informations relationnelles.....	162
1.5. Une dichotomie entre système à base de connaissances formel et les documents..	162
2 - Le document structuré comme support à la capitalisation et la restitution.....	162
2.1. Apports du document numérique structuré dans la capitalisation .....	163
2.2. Apports du document numérique structuré dans la restitution .....	166
Partie IV - Proposition : une approche centrée sur le concept de document.....	168
Chapitre 10 - Approche centrée sur le document structuré.....	169
1 - Document structuré et capitalisation : l'écriture.....	169
1.1. Le document structuré comme espace de conception du langage.....	169
1.2. Le document structuré comme support et mode d'écriture dans la capitalisation : retour d'expérience.....	172
1.3. Infrastructure logiciel de formalisation .....	175
1.4. L'autonomie de l'analyste en question .....	180
2 - Document et restitution.....	180
2.1. Vision globale du processus de restitution .....	181
2.2. Méthode de construction d'une restitution documentaire .....	183
2.3. Définition d'un modèle de domaine.....	186
2.4. Définition d'un modèle de perception.....	189

---

2.5. Modèle d'architecte.....	193
2.6. Construction du document.....	203
3 - Vision intégrée de la capitalisation et de la restitution sur la base de documents structurés .....	211
Chapitre 11 - Développement et mise en œuvre.....	214
1 - Vision globale de la mise en œuvre.....	214
1.1. La capitalisation au travers de documents structurés .....	214
1.2. La restitution .....	222
2 - Réalisation : Les guides d'aide au diagnostic .....	228
2.1. La capitalisation .....	228
2.2. La restitution .....	232
2.3. Retours d'expérience .....	239
3 - Méthodes et outils d'aide à la gestion des actifs de production .....	243
3.1. En capitalisation.....	244
3.2. En restitution.....	251
3.3. Retour d'expérience.....	259
Chapitre 12 - Discussion et Perspectives .....	262
1 - Discussion .....	262
1.1. La construction des documents .....	262
1.2. L'instrumentation de la démarche.....	269
1.3. La pérennité des informations inscrites et de l'instrumentation .....	270
1.4. L'autonomie de l'analyste face à la restitution et la construction d'un document..	271
2 - Perspectives .....	272
2.1. Un approfondissement sur les moyens d'interaction .....	273
2.2. Définition d'un atelier sémiotique et d'une méthode associée pour instrumenter la gestion des connaissances.....	273
Conclusion.....	278
Partie V - Annexes.....	291

## Introduction

Depuis maintenant une quinzaine d'année, entreprises et organisations ont largement pris conscience de l'enjeu crucial que constitue la pérennisation des savoirs et savoirs faire liés à leurs activités. Ces connaissances doivent être considérées comme un capital immatériel qui prend une place de plus en plus importante dans un contexte où l'information, sa gestion, sa diffusion et sa pérennisation jouent un rôle fondamental dans le développement de l'entreprise.

De ce constat, on a alors assisté à l'émergence et au développement de la « Gestion des Connaissances » (Ermine 2003) (Grunstein et Barthès 1996) qui a donné lieu à la conception de méthodes et d'outils (Dieng-Kuntz et al. 2001) et a débouché sur de multiples produits (livres de connaissances (Ermine 2003), mémoires d'entreprise (Grunstein et Barthès 1996), mémoires de projet (Matta 2004)).

Nous ne nous risquons pas ici à définir la connaissance tant il semble qu'il s'agit d'une notion diffuse. Toutefois, nous retenons qu'elle ne peut être abordée comme un objet manipulable (Dieng-Kuntz et al. 2001), même si « *toute connaissances procède d'une inscription matérielle dont elle est l'interprétation* » (Bachimont et Crozat 2004). Ainsi, les produits de l'ingénierie des connaissances ne constituent pas la connaissance en soi, mais bien une représentation de celle-ci (une inscription) portée sur un support (matériel ou immatériel). Ils peuvent se matérialiser sous la forme de documents, papiers ou numériques, ou de systèmes informatiques.

Dans le cadre de cette thèse, notre étude *porte sur les conditions de réalisation de l'inscription d'une connaissance sur un support*. Nous avons ainsi mené une étude empirique dans un contexte d'activité de gestion des connaissances en milieu industriel au sein d'EDF Recherche et Développement. Cette étude a pris pour cadre d'analyse le concept de document (Pédaque 2006) au travers des 3 dimensions qui le caractérise : medium, signe, forme.

Ces trois dimensions ont permis d'orienter l'étude afin d'observer le contexte social dans lequel s'exécute les projets de gestion de connaissances (dimension medium), la construction de sens entre les acteurs (dimension signe) ainsi que la forme des informations inscrites (dimension forme). Dans ces trois axes, nous avons attaché beaucoup d'importance à identifier les outils et supports exploités pour inscrire et faire transiter les informations. Nous avons ainsi bien identifié la prégnance et l'importance de la forme documentaire comme

vecteur de traitement et de diffusion de l'information. On produit ainsi des documents de capitalisation, rédigés sur la base de connaissances identifiées, par exemple, au travers d'entretiens avec des experts. D'autres documents s'attachent à expliquer des résultats d'études. Certains viennent compléter un système informatique, en proposant un discours argumenté en relation avec les résultats produits par ce système : dans ce cas, on est alors dans une situation où les interfaces proposées sont parfois trop complexes pour être appréhendées par un utilisateur n'ayant pas d'expertise sur le système en question. Dans ces différents cas, on peut observer que la construction des documents pourrait très largement être accompagnée. Il s'agirait alors de produire une information plus adaptée à un traitement informatique ou mettant mieux en synergie document et système informatique, représentation narrative adaptée à l'humain et représentation formelle adaptée à la machine.

Ceci nous a conduit à développer une approche méthodologique et instrumentale de construction et de restitutions de documents structurés. Cette approche a été développée en s'inspirant des besoins liés à notre terrain d'étude, mais pourrait s'insérer dans un champ beaucoup plus élargi. Elle trouve ses fondements dans *les champs disciplinaires relatifs au document*, mais également dans la *sémiologie graphique*. La sémiologie graphique, dans notre proposition, a dû être étendue afin de répondre à la construction d'un document. Elle a également fait l'objet d'un ajustement pour prendre en compte les infrastructures logicielles existantes dans les entreprises. En effet, notre volonté n'est pas de proposer une solution « tout-en-un » venant en substitution de l'existant mais de bien de pouvoir s'appuyer sur les outils actuellement déployés et utilisés, malgré les difficultés que cela puisse poser. Ce point introduit une contrainte puisque, contrairement à la sémiologie graphique, il n'est plus uniquement possible de raisonner sur des variables visuelles élémentaires.

L'approche que nous proposons a été développée de manière empirique en s'insérant dans deux projets de gestion des connaissances réels. Le premier, correspondant au développement d'une ressource documentaire relative aux connaissances en maintenance sur de grands ouvrages de production (les guides de diagnostic), nous a permis d'instrumenter l'ensemble de l'activité de gestion d'information allant de son inscription à sa visualisation. Les développements effectués dans ce projet ont fait l'objet d'un déploiement réel sur le terrain. Dans le second projet, en relation avec le développement de Méthodes et Outils d'Aide à la Gestion des Actifs de Production, nous avons effectué des propositions, sur la base de notre approche, pour appuyer une méthode de capitalisation de connaissances et compléter un système à base de connaissances existant, en particulier sur le plan de la restitution.

Cette thèse est structurée en quatre parties.

La Partie I - (page 20) définit notre problématique. Tout d'abord, elle décrit notre terrain d'étude, en introduisant la gestion des connaissances à un niveau général, puis se centre sur le cas particulier du milieu industriel à travers l'exemple de EDF.

La Partie II - (page 62) doit permettre de situer notre approche théorique qui articule 3 champs disciplinaire : l'ingénierie documentaire (Chapitre 3 - page 63), l'ingénierie des interfaces homme machine (Chapitre 4 - page 83) et ce que nous nommons l'ingénierie de la forme (Chapitre 5 - page 102) dans lequel nous regroupons visualisation d'information et sémiologie graphique.

La Partie III - (page 127) décrit l'analyse empirique des systèmes de gestion de connaissances que nous avons mené. Elle est conduite avec comme cadre d'analyse le document. Ainsi, nous avons analysé les systèmes de gestion de connaissances mis en œuvre selon 3 dimensions : medium (Chapitre 6 - page 130), signe (Chapitre 7 - page 137), forme (Chapitre 8 - page 148). Cette analyse a porté sur l'activité du groupe de gestion des connaissances dans lequel nous avons évolué. Elle est également ancrée dans deux projets réels sur lesquels nous avons à la fois observé et instrumenté capitalisation et restitution des connaissances. Ceci nous a permis de caractériser les systèmes de gestion de connaissances mis en jeu, leur fonctionnement ainsi que leurs limites afin d'en déduire les besoins complémentaires qui en découlent (Chapitre 9 - page 160).

Dans la Partie IV - (page 168), nous proposons une approche méthodologique et instrumentale de construction et de restitution de documents structurés (Chapitre 10 - page 169). Le document structuré est alors proposé comme une réponse à la capitalisation et à la restitution de connaissances. Cette approche a été confrontée aux deux projets réels ayant supportés notre étude (Chapitre 11 - page 214), ce qui a permis d'obtenir un retour d'expérience.

Enfin, avant de conclure, nous discutons notre approche, en dehors de nos cas d'application, pour en identifier les apports et les limites, et tracer quelques perspectives (Chapitre 12 - page 262).

Ce mémoire propose également deux annexes. La première (Annexe 1 - page 293) propose un complément d'analyse relative aux moyens de visualisation d'informations et d'interaction sur différents logiciels ou formats informatique venant en complément du Chapitre 10 -2.5.b page 196. La seconde (Annexe 2 - page 305) correspond a un extrait d'un rapport produit

pour EDF concernant le travail réalisé sur un de nos cas d'application. Il s'agit d'une description plus précise de la réalisation produite.

# Partie I - Problématique

# Chapitre 1 - Le terrain

## 1 - Gestion des connaissances

Définir la Gestion des connaissances se révèle un exercice difficile si l'on considère l'étendue du concept. Il se situe à l'intersection de plusieurs champs disciplinaire tels que l'économie et la gestion, les sciences cognitives ou encore l'informatique avancée et plus largement les technologies de l'information et de la communication (TIC).

La gestion des connaissances peut être désigné comme « *un ensemble de concepts et d'outils permettant aux membres d'une organisation de travailler ensemble et de faire le lien entre informations disponibles, production de connaissances et développement des compétences individuelles, collectives et organisationnelles* » (Ermine 2000). Cette définition est assez générale étant donné le caractère pluridisciplinaire de cette activité qui peut être conduite selon :

- La dimension organisationnelle et communautaire,
- La dimension modélisation de l'activité et de l'information,
- La dimension technologique.

### 1.1. La gestion des connaissances à la croisée de 3 dimensions

La gestion des connaissances peut ainsi être conduite par la *dimension organisationnelle et communautaire*. Ces études doivent permettre de développer des démarches managériales et de gestion pouvant accompagner le pilotage d'actifs de l'entreprise et conduire le changement. Cette orientation peut être illustrée au travers des travaux relatifs aux communautés de pratique (Wenger 2001). Nous évoquerons également les travaux de Nonaka et Takeuchi (Nonaka et Takeuchi 1995). Ils proposent le modèle SECI (Socialization – Externalization – Combination – Internalization) qui introduit une conceptualisation du processus de création de connaissances. De nombreux autres travaux peuvent être associés à cette direction pour l'analyse de réseaux de santé (Grenier et Pauget 2006) (Benard 2007) ou encore l'analyse de l'entraide en entreprise (Delalande 2007).

La gestion des connaissances peut également être guidée par la *modélisation de l'activité et de l'information*. Dans cette orientation, la démarche doit permettre d'aboutir à une

représentation des connaissances pouvant être partagée, transférée, valorisée et pérennisée. La connaissance n'est pas en soi un « objet manipulable » (Dieng-Kuntz et al. 2001). Ainsi, nous sommes peu ancrés dans les paradigmes de connaissances tacites et explicites par exemple. N'appréhendant pas la connaissance comme un objet formalisable mais plutôt comme un élément s'ancrant dans un processus cognitif, la notion de connaissance explicite semble plutôt se rapprocher d'une information inscrite sur un support dans notre positionnement. C'est alors dans le processus d'interprétation de cette inscription que peut naître une connaissance qui sera assimilée par une personne. Le processus d'interprétation reste toutefois indéterministe au sens où chaque personne va interpréter les choses selon ses spécificités liées à son origine ou encore sa culture. Il s'agit donc d'accompagner cette interprétation pour tenter de la guider au mieux. Aussi, dans l'étude que nous proposons dans ce mémoire, nous ne traitons que de représentations de connaissances. L'enjeu étant de clarifier au mieux le langage de représentation utilisé dans sa construction pour en faciliter l'interprétation. Nous appréhendons alors une représentation de connaissance comme une inscription dans un langage particulier sur un support : cette représentation constitue alors une information. C'est en particulier à ce niveau que se situe l'ingénierie des connaissances relative au développement d'artefacts permettant de supporter l'activité de gestion des connaissances. Nous y reviendrons dans la section suivante.

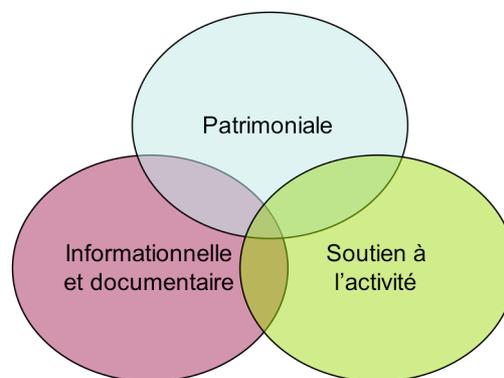


Figure 1 Différentes vues relatives à l'activité de représentation de connaissances

Il est possible de dégager plusieurs vues relatives à cette activité de représentation de connaissances. La première est *informationnelle et documentaire*. Dans ce premier cas, l'objectif est de faciliter l'accès à des informations. Il s'agit de traiter les problématiques liées à l'hétérogénéité des informations en termes de forme et de structuration, de volume ainsi que d'évolution. Ce champ d'activité qui historiquement traitait de la gestion des livres et des documents représentait le terrain privilégié des sciences de l'information (Le\_Coadic 1994) et

notamment la bibliothéconomie et de l'archivistique (Gagnon-Arguin 1992; Mas 2004). Par la suite, il s'est étendu à l'informatique avec la numérisation des supports informationnels.

Une deuxième vue correspond à la *vue patrimoniale*. C'est une vision largement emmenée par la vague du « Knowledge Management » des années 90 et la prise de conscience que les connaissances relatives aux savoirs et savoirs-faire d'une entreprise ou d'une organisation représentent un capital immatériel de premier plan. L'enjeu est ici de conserver ces connaissances afin de constituer une histoire technique et organisationnelle. Les livres de connaissances (Ermine 1996; Ermine et al. 1997) peuvent être associés à cette vue ainsi que les approches relatives aux bases de récits d'expérience ou storytelling (Soulier 2003).

Enfin, une troisième vue correspond au *soutien à l'activité*. La modélisation de l'activité ou de l'information peut fournir une aide à la résolution de problème. C'est ici que se positionnent les travaux issus de l'intelligence artificielle par exemple. Elle peut également fournir une aide à la gestion de l'activité en contexte pour un individu. La représentation des connaissances sera plus ou moins codifiée et manipulable par la machine, selon que l'on s'inscrit dans une approche calculatoire et computationnelle ou non.

Quelle soit menée selon une approche organisationnelle ou selon une approche de modélisation, la gestion des connaissances intègre nécessairement *une dimension technologique*. En effet, si la cible correspond au développement d'une ressource papier, tout document de ce type prend une forme numérique à un moment ou à un autre de son cycle de vie. De même, les systèmes informatiques développés dans une approche de soutien à l'activité qu'ils soient appelés « systèmes de gestion de connaissances » ou « systèmes à base de connaissances », se concentrent sur le traitement et la valorisation d'informations numérisées. Enfin, les systèmes organisationnels sont de plus en plus souvent supportés par des infrastructures numériques (email, téléphone, forum, wiki...) qui médiatisent les échanges entre les personnes. L'information numérique devient incontournable et prend alors une place importante, voire cruciale.

Ainsi, la gestion des connaissances, s'appuyant sur les sciences et technologies de l'information, doit proposer des solutions pour gérer et valoriser au mieux les informations existantes, qu'elles soient numériques ou non. La gestion des connaissances s'inscrit également en créatrice d'informations et d'outils, s'agissant tout aussi bien de documents que de systèmes informatiques. Dans cette perspective, de nombreuses démarches ont été mises en oeuvre pour accompagner l'activité de gestion des connaissances aboutissant souvent à la production de nouvelles ressources informationnelles. La section suivante évoque sans

volonté d'être exhaustive, quelques méthodes développées pour mener une démarche de gestion des connaissances.

## 1.2. Quelques méthodes

La définition proposée par Ermine de la gestion des connaissances, permet d'identifier deux dimensions inhérentes à l'activité de gestion des connaissances. L'une renvoie à l'aspect méthodologique et l'autre au développement d'outils supportant cette activité. Il n'y a bien évidemment pas une dichotomie entre ces 2 dimensions, l'outil impactant sur la méthode et inversement. L'engouement et l'enjeu qui gravitent autour de la gestion des connaissances ont donné lieu au développement de nombreuses méthodologies ayant pour ambition de supporter cette activité en se concentrant largement sur la problématique de l'acquisition ou capitalisation des connaissances.

Les méthodes développées sont donc nombreuses (Rose Dieng et al. 98; Dieng-Kuntz et al. 2001) et nous n'en citons que quelques unes ici. Tout d'abord, nous évoquons les méthodes comme KADS ou KOD (Dieng-Kuntz et al. 2001). Elles se concentrent sur le développement de ce type de systèmes à base de connaissances permettant la manipulation d'une représentation des connaissances par l'ordinateur. Ces méthodes se situent plus dans le prolongement et la logique computationnelle des travaux d'intelligence artificielle se focalisant largement sur le soutien à l'activité et mettant en œuvre des algorithmes de simulation et de raisonnement par exemple.

Dans une orientation plus centrée sur une vue informationnelle, la méthode REX (Malvache et Prieur 1993) propose de capitaliser des éléments d'expériences sous la forme de « fiches de connaissances ». L'ensemble de ces fiches de connaissances doit alors constituer une mémoire. La méthode MKSM (Methodology for Knowledge System Management) pour sa part introduit des phases de cadrage, de modélisation et de capitalisation permettant le développement de diagrammes ayant pour ambition de décrire la connaissance selon différents points de vue. L'ensemble de ces diagrammes permet d'aboutir à un livre de connaissances (Ermine 1996; Ermine et al. 1997). A notre sens, fiches de connaissances et livres de connaissances ne constituent, non la connaissance en elle-même, mais bien une représentation de celle-ci inscrite sur un support. Cette inscription qui peut alors être consultée, sera, en effet, interprétée selon les propriétés de chaque individu. Cela justifie l'importance de la sémantique relative à cette inscription qui doit être identifiée et précisée afin d'accompagner l'appropriation.

Dans le prolongement de la méthode MKSM, la méthode MASK (Ermine 2003) ne se limite pas à la problématique de capitalisation mais veut couvrir l'ensemble du processus de gestion des connaissances se composant de la capitalisation, de l'appropriation et de la pérennisation des connaissances capitalisées. La capitalisation ne présente jamais un caractère exhaustif puisque les démarches mises en œuvre ne permettent d'identifier qu'une partie des connaissances de l'entreprise notamment celles pouvant être représentées et inscrites sur un support : il ne s'agit donc que d'un échantillonnage. On introduit alors souvent la notion de *connaissances cruciales* (Grunstein et Barthès 1996) comme un sous-ensemble d'informations qui sont d'une grande importance pour l'entreprise. Beaucoup d'études portent alors à la fois sur le choix des connaissances à représenter (Saad et al. 2005), mais également sur les techniques d'entretien d'experts permettant d'identifier les connaissances qu'ils possèdent tout en stimulant leur créativité (Bouzaiène et al. 2002; Bouzaiène 2005). Les diagrammes collationnés sous la forme d'un livre de connaissances dans la méthode MKSM représentent un guide au développement de modèles conceptuels pouvant appuyer le développement d'outils tels les systèmes à base de connaissances (Matta 2004).

Les quelques méthodes évoquées permettant de produire de nouvelles ressources informationnelles ou de nouveaux outils. Ils constituent autant d'instruments pour accompagner les activités métier. Ces ressources sont plus ou moins numérisées, plus ou moins structurées et codifiées offrant ainsi un éventail de possibilités de valorisation de l'information. Cette démarche de développement d'outils renvoie à ce que l'on peut nommer *l'Ingénierie des connaissances*.

## 2 - Instrumenter la Gestion des Connaissances : l'Ingénierie des Connaissances

La gestion des connaissances a donné lieu à une longue tradition de création d'artefacts plus ou moins complexes, plus ou moins techniques, et plus ou moins automatisés. Ils avaient pour ambition de supporter capitalisation, représentation, diffusion, et pérennisation d'une représentation des connaissances. Cette tradition trouve, pour partie, ses origines dans le domaine de l'intelligence artificielle, l'archétype étant le développement des systèmes expert (Hatchuel et Weil 1992).

## 2.1. Des systèmes experts à l'orientation dans les supports numériques

Les systèmes experts étaient définis comme visant à « à piéger la connaissances naturelle détenue par des experts humains sous la forme la plus proche de son expression par l'expert ; celle-ci captée, on pourra imiter les raisonnements de certains individus pour résoudre informatiquement certains problèmes » (Hatchuel et Weil 1992). Le point de vue qui prévalait ici était la conception d'un système qui simulerait le « raisonnement de l'expert », la problématique étant de définir le bon niveau d'expression et la forme de représentation adéquate, pour la connaissance. Le bon niveau d'expression et la bonne forme sont en grande partie liés aux possibilités d'un moteur d'inférence : les connaissances modélisées doivent être manipulables par celui-ci. Les limites de ce type d'approche sont aujourd'hui connues notamment parce qu'elles appréhendaient la connaissance comme un objet formalisable dans un cadre théorique et implémentable permettant notamment d'intégrer des mécanismes d'inférence (Simon 1978; Newell 1981). Les systèmes étaient basés sur une logique calculatoire, la limite étant nécessairement que tout résultat quantitatif, aussi pertinent soit il, ne peut faire l'économie d'une interprétation : le résultat d'un calcul renvoie à la démonstration et doit, lors de sa diffusion, faire l'objet d'une explication, c'est à dire une argumentation.

A la suite de ces travaux plutôt centrés sur la résolution automatique de problèmes, on a pu observer une évolution vers des approches majoritairement centrées sur l'assistance utilisateur. Le système expert n'était plus un expert omniscient mais un assistant coopératif « intelligent » (Cawsey 1995). Ce positionnement a notamment guidé les travaux de recherche en explication (Baker 1996; Baker 2000). Le système ne renvoie plus à un résultat mais argumente celui-ci afin de préciser « le pourquoi et le comment » (Cawsey 1994). On ne se situe plus dans la démonstration, mais dans l'argumentation, l'outil de simulation informatique devenant outil d'aide à la résolution de problème. Le résultat n'est plus une vérité instituée, car il n'est plus toujours considéré comme efficace, mais pourra être accompagné de traces d'exécution de l'outil devant expliciter le résultat ou encore de connaissances complémentaires extérieures au système. Le résultat peut être soumis à la négociation sous la forme d'une interaction entre utilisateur et machine (Baker 1994; Baker 2000).

Ceci nous amène à une vue alternative à l'IA telle qu'elle a pu être fondée initialement. Ce positionnement également défendus dans le cadre des conférences *Cooperative System Design*

---

(COOP) se centrait sur le développement de systèmes permettant une forme de « coopération » entre le système et l'utilisateur. Les approches sont alors d'avantage inspirées du Computer Supported Cooperative Work (CSCW) (Schmidt et Simone 1996; Ackerman 1997; Dourish 1999) et des Sciences de l'Information. Il ne s'agit plus de se focaliser exclusivement sur un niveau d'expression propice à une manipulation par la machine, mais bien de remettre en avant la dimension sociale et la dimension organisationnelle. Ce point de vue privilégie la diffusion, l'appropriation et le partage d'une représentation des connaissances. Les mémoires d'entreprises<sup>1</sup> constituent un exemple produit avec cette philosophie. Le point de vue est bien celui d'un support à l'activité même si le but des mémoires d'entreprise peut également être présenté comme devant fournir « *la bonne connaissance ou information, à la bonne personne, au bon moment et au bon niveau* » pour permettre la décision. (Grunstein et Barthès 1996). L'aspect support à la décision apparaît ici, de même que la problématique de recherche d'information qui devient de plus en plus complexe du fait de la profusion des supports numériques.

L'ingénierie des connaissances s'appuie donc à la fois largement sur les techniques d'IA (Boughzala 2001) mais également sur des champs disciplinaires plus proches des sciences humaines et sociales. Elle vient en prolongement de la gestion des connaissances telle que définie précédemment et peut être présentée comme « *une manière plus sophistiquée que la transcription des connaissances pour capter une parcelle de connaissances tacites* » (Boughzala et Ermine 2003).

Cette « manière plus sophistiquée » correspond à la conception et au développement d'instruments et d'outils pouvant à la fois supporter la gestion des connaissances, mais également faire face à la prolifération et à l'expansion de plus en plus grande de l'information numérique. En ce sens, l'Ingénierie des Connaissances doit « *faciliter l'explicitation de connaissances tacites possédées par des informateurs particuliers, les experts, ou déposées dans des supports informationnels de manière implicite, des textes ou des bases de données de divers formats* » (Zacklad 2000). Cette définition intègre non seulement la dimension humaine de la gestion des connaissances relative au fait que les connaissances sont dispersées au travers d'individus ; la psychologie, l'ergonomie, la sociologie ou encore l'IA ont largement étudié cet aspect pour tenter d'acquérir, de capitaliser, une représentation de ces connaissances dispersées (Charlet 2003). Elle intègre également la dimension technique de la

---

<sup>1</sup> « le terme "mémoire d'entreprise" désigne l'ensemble des savoirs et savoirs-faire en action, mobilisés par les employés d'une entreprise pour lui permettre d'atteindre ses objectifs (produire des biens ou des services) » Jean-Paul Barthès, Rose Dieng et G. Kassel (1999). "Mémoire d'Entreprise." bulletin de l'AFIA(n°36).

gestion des connaissances et l'enjeu crucial des supports informationnels, qu'ils soient papier ou numérique, dont le volume est en constante croissance.

Ainsi, l'objectif de l'ingénierie des connaissances est « *de concevoir des instruments et outils de la connaissance s'intégrant au travail intellectuel, qu'il soit individuel ou collectif, cognitif ou social* » (Bachimont 2004). Bachimont pose d'emblée que « *le problème posé par une telle discipline est la diversité des savoirs qu'elle doit mobiliser, et l'étendue de ses objets d'étude: toute l'activité humaine est en effet potentiellement concernée par un tel programme* » (Bachimont 2004). Le champ d'exploration est donc immense, aussi nous nous focaliserons sur un aspect : l'étude portée dans ce rapport s'intéresse en particulier aux informations numériques pouvant être valorisées ou résultant des activités de gestion des connaissances. Il s'agit tout aussi bien de s'attarder sur les modalités de création de ces ressources informationnelles que sur la valorisation de l'existant. Au-delà, de nombreux fonds documentaires non numérisés, ou dont la numérisation se limite à un plan de classification et à une description des documents selon des attributs standards, les ressources numériques sont extrêmement importantes en volume. Dans la section suivante, nous proposons de nous attarder sur ces supports numériques.

## 2.2. L'avènement des supports numériques

Les supports numériques d'information envahissent de plus en plus nos pratiques. Nous sommes en permanence confrontés à des documents, des interfaces, des systèmes et des bases de données que nous consultons ou pour lesquelles nous sommes fournisseurs d'information. Le point de départ de nombre d'activités est ainsi d'identifier et de réunir les informations pour ensuite les analyser, les traiter et les valoriser. Il s'agit de les croiser, de les mettre en relation, de les comparer, etc... La difficulté première est à la fois le fait que l'on a pas toujours connaissance des informations nécessaires à la tâche, ni de celles disponibles. De plus, il peut être difficile de s'orienter dans un grand volume d'information. Ainsi, Bachimont introduit l'Ingénierie des connaissances comme un moyen de « *pallier à la désorientation induite par l'avènement des supports numériques* » (Bachimont 2004). Nous allons évoquer ici successivement trois formes d'informations couramment portées sur des supports numériques :

- Les documents
- Les données structurées

- Les systèmes informatiques

### *2.2.a Les documents*

Le document est abordé ici sans en préciser une réelle définition, qui sera introduite dans la revue de littérature. Nous traitons ici de tout type de documents, qu'il soit numérique ou papier. En premier lieu, précisons que de nos jours quasiment tout document même si il est diffusé sur papier prend une forme numérique à un instant donné (Crosnier 2005). Dans les entreprises, les auteurs de documents se sont profondément diversifiés puisque le déploiement de la micro-informatique à tous les personnels a profondément changé, par exemple, le travail des secrétaires qui supportent majoritairement les activités de gestion, plutôt qu'une tâche de claviste par exemple. Tout le personnel est devenu potentiellement auteur de documents qui transitent de plus en plus vite dans l'entreprise au travers des systèmes d'information. L'intégration de systèmes de gestion électronique de documents appuyés par les technologies Lotus notes ou maintenant XML (eXtended Markup Language) associées au processus qualité de gestion de projet tend à en augmenter la production. Rapports d'Etude, Cahiers des charges ou encore Compte-Rendu de réunions médiatisent une image de l'évolution et l'avancement des projets, visualisable dans les systèmes de gestion électronique de documents (GED), à mesure de leur changement de statut (en cours de rédaction, pré-diffusion, diffusion).

Le déploiement de l'informatique bureautique introduit ainsi une facilité dans la production de document. Elle est également liée à l'appropriation des logiciels Wysiwyg<sup>2</sup> (« What you see is what you get ») et en premier lieu à l'avènement des traitements de texte tel que Microsoft Word. Tout un chacun est ainsi en mesure de produire un document avec un rendu, ce qui était autrefois réservé aux professionnels de l'édition. Les éditeurs de document permettent ainsi d'associer différentes modalités de représentation (graphique, textuelle). Chacun peut produire ses documents, de façon autonome, selon ses préférences. De plus, on assiste à un vaste usage du « copier – coller » qui permet à la suite de quelques manipulations de fusionner, reconstruire, recomposer le document à souhait.

### *2.2.b Données structurées*

Nous définissons les données structurées comme un ensemble de symboles qui sont organisés logiquement selon des règles particulières, identifiables et uniformes pour un même jeu de données. Les fichiers textes délimités (CSV : Comma-separated values) constituent un

---

<sup>2</sup> Les logiciels Wysiwyg sont caractérisés par le fait que la visualisation qu'ils présentent correspond au rendu que l'on obtiendra sur écran ou sur papier.

exemple caractéristique. Ils représentent des données tabulaires dont on peut voir une illustration en Figure 2. Ce type de fichier se compose le plus souvent d'une première ligne qui identifie (indexe) les données présente sur les lignes suivantes. Chaque donnée est alors délimitée par un caractère spécifiant le passage sur une ligne du fichier, d'une donnée à une autre. Les caractères exploités sont classiquement l'espace, la tabulation ou encore le point-virgule.

Sexe, Prénom, Année de naissance	Sexe	Prénom	Année de naissance
M,Alphonse,1932	M	Alphonse	1932
F,Béatrice,1964	F	Béatrice	1964
F,Charlotte,1970	F	Charlotte	1970

Figure 2 Exemple de fichier CSV dont les informations sont délimitées par « , » et en regard son équivalent matriciel (source : Wikipedia)

La structuration logique des fichiers CSV est devenue d'un usage courant, ce qui en fait un outil intégré à de nombreuses pratiques. En particulier, le CSV supportent les fonctions d'importation et d'exportation de données facilitant ainsi l'échange et la migration d'importants volumes de données entre systèmes informatiques. Au regard de l'échange de données et dans un contexte industriel, ce type de format est largement exploité dans le cadre de la remontée de données de mesure par exemple. Ceci s'explique par le fait que même des équipements de « bas niveau », avec une électronique peu complexe, peuvent enregistrer les données dans ce type de format simple. Une difficulté au regard de ce format reste que la structuration des données est limitée à deux dimensions. De plus, la définition de l'organisation tabulaire peut varier si une norme ou des conventions ne sont pas mises en place.

Les données formatées en XML (eXtended Markup Language) représentent une alternative au format CSV. Il s'agit d'un langage informatique de balisage basé sur une simplification de la norme SGML (Standard Generalized Markup Language). Ce langage, dont on peut voir un aperçu sur la Figure 3, est de plus en plus exploité. Il permet à la fois d'identifier les informations contenues dans un fichier par un balisage prenant la forme « <nom\_balise> ».

Ces balises peuvent être imbriquées de manière arborescente permettant de structurer l'information sur plusieurs niveaux.

```
<?xml version="1.0" encoding="UTF-8"?>
  <!-- Commentaire -->
  <élément-document xmlns="http://exemple.org/" xml:lang="fr">
    <élément>Texte</élément>
    <élément>élément répété</élément>
    <élément>
      <élément>Hiérarchie récursive</élément>
    </élément>
    <élément>Texte avec <élément>élément</élément> mêlé</élément>
    <élément/><!-- élément vide -->
    <élément attribut="valeur"></élément>
  </élément-document>
```

Figure 3 Exemple de fichier XML (source : Wikipedia)

L'eXtended Markup Language se veut un format générique et véhicule avec lui un ensemble d'outils qui permettent de manipuler les contenus structurés indépendamment de leur origine et de valider leur bonne structuration en s'appuyant sur des modèles de documents. L'XML s'impose de plus en plus comme le format dédié à l'échange de données informatisées (EDI). Cela est à la fois dû à sa grande souplesse mais également au fait qu'il constitue un standard accompagné d'un ensemble d'outils permettant de facilement le manipuler. Une forme d'utilisation de ce standard correspond aux flux RSS (cf. Figure 4). Il s'agit de flux de données structurées qui peuvent être lues de manière standard par des applications telles qu'un navigateur web ou un lecteur spécialisé. Ces flux sont de plus en plus répandus avec le développement du Web 2.0 car ils sont exploités dans les mécanismes de syndication<sup>3</sup>.

```
<rss version="2.0">
  <channel>
    <title>Mon site</title>
    <description>Ceci est un exemple de flux RSS</description>
```

<sup>3</sup> La syndication renvoie à l'abonnement à un flux de données permettant d'obtenir régulièrement les mises à jour.

```
<lastBuildDate>Wed, 27 Jul 2005 00:30:30 -0700</lastBuildDate>
<link>http://www.example.org</link>
<item>
  <title>Actualité N°1</title>
  <description>Ceci est ma première actualité</description>
  <pubDate>Mon, 25 Jul 2005 00:30:30 -0700</pubDate>
  <link>http://www.example.org/actu1</link>
</item>
<item>
  <title>Actualité N°2</title>
  <description>Ceci est ma deuxième actualité</description>
  <pubDate>Tue, 19 Jul 2005 04:32:51 -0700</pubDate>
  <link>http://www.example.org/actu2</link>
</item>
</channel>
</rss>
```

Figure 4 Un exemple de flux RSS (source : Wikipedia)

### 2.2.c Systèmes informatiques

Les systèmes informatiques dans leur globalité peuvent être vus comme une forme d'information à part entière. Bien évidemment ces systèmes s'appuient sur des données structurées telles que nous venons de l'évoquer. On peut d'ailleurs citer les systèmes de gestion de base de données qui constituent des systèmes dédiés à l'organisation de données à la fois très structurées et très typées. Cette organisation est le plus souvent régie par un modèle relationnel correspondant à une réelle grille de lecture de la base. Nous pensons mettre l'accent sur deux autres formes d'information contenues dans les systèmes informatiques :

- Les algorithmes
- Les interfaces homme machine

Tout d'abord, « *un algorithme est un moyen pour un humain de présenter la résolution par calcul, d'un problème à une autre personne physique (un autre humain) ou virtuelle (un*

*calculateur*) » (*Wikipedia*). La définition permet dans les systèmes informatiques de programmer la logique de traitement sur les données sous la forme d'un langage de programmation. Les algorithmes constituent donc une forme de représentation d'information essentielle puisque à l'heure où les supports numériques se multiplient, de même que les traitements numériques de leur contenu, ils prennent une importance énorme. Ils véhiculent un processus de calcul pouvant être appréhendé comme une opération de réécriture des données selon une logique et des intentions particulières. L'algorithme est souvent dilué dans le code source ce qui en fait une information qui n'est pas nécessairement explicite : le fonctionnement du système, la logique de calcul, reste alors tacite pour les utilisateurs.

Les interfaces homme machine, quant à elles, représentent le pont entre l'utilisateur et le système informatique. Elles permettent d'offrir une visualisation des données abstraites contenues en mémoire (sur les supports numériques) et d'interagir avec elles. L'interface va donc traduire les données en mémoire sous une forme d'inscription qui lui est propre. Au-delà de cette forme de représentation, elle véhicule à l'image de l'algorithme une intention. Sa construction, par exemple l'organisation des données qu'elle permet de visualiser, contient une part d'information qui une fois encore n'est pas toujours explicite.

### 2.3. Les produits de l'ingénierie des connaissances

Si l'ingénierie des connaissances tente de pallier la désorientation dans les supports numériques tout en valorisant au mieux les informations, elle constitue également une activité qui produit de nouvelles ressources informationnelles. Nous considérons trois types de ressources produites :

- Des documents,
- Des systèmes à base de connaissances,
- Des systèmes d'organisation de connaissances.

#### 2.3.a Des documents

Le *document* est un vecteur de diffusion de l'information omniprésent dans les activités humaines et en particulier celles des « travailleurs du savoir » (« knowledge workers », Peter Drucker dans (Alberts et Bertrand-Gastaldy 2006)). L'activité des travailleurs du savoir est caractérisée par le fait que la tâche intellectuelle domine sur le travail manuel (Alberts et Bertrand-Gastaldy 2006). Lorsque nous évoquons le document ici, il s'agit d'un document

écrit pouvant également contenir d'autres modalités d'information telles que, par exemple, des schémas ou des photos. Sa numérisation a ouvert un vaste champ de possibilités le rendant interactif et dynamique : il peut alors contenir animation, sons et vidéos. Les livres de connaissances (Ermine 2003) peuvent se matérialiser sous la forme de documents tout comme les mémoires d'entreprise ou encore les mémoires de projet.

Les *mémoires de projet* (Matta 2004) sont développées pour conserver une trace des connaissances dans les processus de conception afin de pouvoir être réinvesties, par exemple. Les *mémoires d'entreprise* (Grunstein et Barthès 1996) sont introduites comme une « *représentation explicite pertinente et désincarnée, des connaissances et informations dans une organisation* ». Une mémoire d'entreprise a pour objet de fournir « *la bonne connaissance ou informations, à la bonne personne, au bon moment et au bon niveau* » pour permettre la prise de décision (Grunstein et Barthès 1996). Notons que cet objectif peut également faire référence au développement des systèmes d'aide à la décision et donc aux systèmes à base de connaissances que nous aborderons par la suite. Dans une vision moins formelle, « *le but d'une construction de mémoire d'entreprise est moins ambitieux qu'un système expert, au lieu de viser la résolution automatique pour une tâche (avec des capacités automatiques de raisonnement), une mémoire d'entreprise doit plutôt aider l'utilisateur, en lui fournissant des informations appropriées de l'entreprise, mais en lui laissant la responsabilité d'une interprétation et d'une évaluation contextuelle de ces informations* ». Kühn et Abecker (1997) cité dans (Dieng-Kuntz et al. 2001).

Si dans cette citation le but est évoqué comme moins ambitieux qu'un système expert, il convient tout de même de construire les mémoires d'entreprises ou plus généralement ce que nous nommerons des documents de capitalisation de connaissances avec précaution. En effet, ces documents revêtent une importance cruciale dans les entreprises puisqu'ils tendent à devenir, si ce n'est des documents de référence, des points d'accès à de nombreuses informations. Dans ce cas, ils peuvent prendre la fonction de *guide de lecture* permettant d'éviter la désorientation évoquée par Bachimont. Aussi, convient il de s'attacher à ce que leur construction sur la forme et le fond accompagne le lecteur vers une interprétation des informations proches de celle de l'auteur. La sémantique des informations qui y sont portées doit alors y être précisée.

### 2.3.b *Systèmes à base de connaissances*

Les systèmes à base de connaissance peuvent être vus comme une évolution des systèmes experts. Si les systèmes experts se fondaient sur le développement d'une représentation du savoir expert sous une forme permettant d'être valorisée au travers d'un système d'inférence, les systèmes à base de connaissances ne font pas une référence exclusive à ce type de savoirs experts formalisés. Il ne s'agit plus uniquement de développer des systèmes à base de règles avec les limites que l'on connaît mais également de mettre à profit d'autres sources d'information. Ces informations peuvent tout aussi bien être des documents, que des bases de données ou encore des sources de données moins structurées. Ainsi, la problématique du développement des systèmes à base de connaissance s'est élargie pouvant tout aussi bien soutenir le développement des hypertextes comme une aide à la navigation par exemple, au raisonnement à base de cas ou encore être investis dans le développement de systèmes coopératifs. Une illustration renvoie à l'analyse d'informations relatives aux compétences d'employés ainsi qu'aux projets auxquels ils ont participé pour favoriser l'entraide en entreprise (Delalonde 2007).

Le fonctionnement d'un système à base de connaissances reste toutefois centré sur l'analyse d'une représentation de connaissances par un algorithme traduisant une stratégie d'exploitation des données. De là, les systèmes à base de connaissances peuvent être vus comme « *des systèmes sémiotiques de manipulation d'inscriptions symboliques, dont le fonctionnement informatique doit permettre à un utilisateur d'interpréter et de comprendre le système dans le cadre de son activité et de ses usages, en utilisant les termes du domaine* » (Charlet & Bachimont, 1998). Dans cette définition, comme pour les documents évoqués précédemment, les notions d'interprétation et de compréhension de la part de l'utilisateur sont évoquées. Là encore, le développement des systèmes à base de connaissances et notamment de leur surface visible à l'utilisateur (ie. leurs interfaces homme machine), doit s'attacher à bien véhiculer la sémantique des informations proposées afin que l'interprétation faite par l'utilisateur ne soit pas erronée.

### 2.3.c *Les systèmes d'organisation de connaissances*

Les systèmes d'organisation de connaissances se différencient des systèmes à base de connaissances au sens où ils ne se focalisent pas sur l'analyse et le traitement d'une représentation de connaissances. Il s'agit plutôt d'apporter une solution à la problématique de désorientation dans un espace informationnel. C'est notamment le champ disciplinaire des

sciences de l'information qui dans le cadre de la gestion des livres et des documents a conçu de nombreux instruments de ce type. Les *classifications bibliographiques* (ex : la classification décimale de Dewey (Béthery 2005)) pour les ouvrages en bibliothèque, ou encore les *plans de classement* (Mas 2004) pour les documents administratifs en sont des exemples. Ces systèmes d'organisation offrent un appui à la recherche documentaire rejoignant les objectifs des documents et des systèmes à base de connaissances évoquées précédemment. La différence est que la démarche, le processus de recherche, est centré sur l'individu. Ainsi, les systèmes d'organisation de connaissances sont un accompagnement dans la réalisation de la tâche, sans avoir l'ambition de fournir instantanément le bon résultat. Ils s'ancrent dans le domaine de la recherche ouverte d'information (L'Hédi Zaher et al. 2007).

Dans le cadre de la recherche d'information, on peut également citer d'autres systèmes d'organisation de connaissances. Tout d'abord, les *taxonomies* qui proposent une organisation hiérarchique d'un vocabulaire contrôlé. La notion de vocabulaire contrôlé renvoie à la définition d'un ensemble de termes définis par un groupe, une communauté, appartenant, par exemple, à un même domaine de compétence. L'organisation hiérarchique introduite par la taxonomie permet, le plus souvent, de définir une spécialisation entre les termes (cf. Figure 5). Les *thesaurus* correspondent à une évolution des taxonomies car ils ne se contentent pas d'offrir la spécialisation entre les termes mais permettent également de spécifier des informations connexes à chaque élément qui constitue la taxonomie. Ils sont largement utilisés en recherche d'informations permettant d'accompagner un utilisateur dans la construction de ses requêtes face à un moteur de recherche. Enfin, nous évoquerons les *réseaux sémantiques* ou *graphes de concepts* qui correspondent à des graphes étiquetés de concepts introduisant des relations de catégorisation (relation « is a ») ainsi que des relations méronymiques (relation « part-of ») spécifiant la composition (relation partie-tout) (Un exemple est le *lexidiom* proposé par la société MEMODATA spécialisé en traitement automatique du langage - <http://www.memodata.com/> ). La structuration des concepts dépasse ici la hiérarchie, impliquant que les réseaux sémantique ne répondent plus totalement, ou plus difficilement, à certaines propriétés formelles : leur exploitation par la machine est alors plus complexe. Les propriétés formelles demeurent une valeur recherchée dans le développement de systèmes d'organisation de connaissances expliquant notamment l'engouement pour le développement des ontologies formelles.

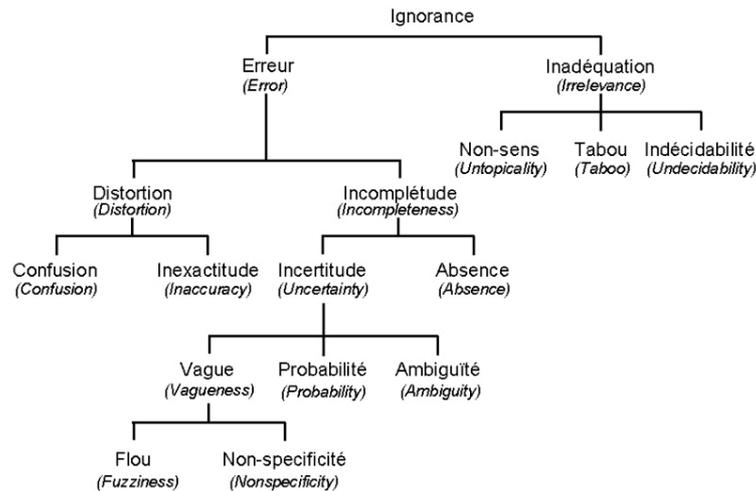


Figure 5 Exemple de taxonomie (source : [www.theses.ulaval.ca/2004/22242/ch02.html](http://www.theses.ulaval.ca/2004/22242/ch02.html) )

Même si le terme a largement été mis en avant ces dernières années en informatique, l'ontologie est une notion qui vient de la philosophie. « *En informatique et en science de l'information, une ontologie est un ensemble structuré de concepts permettant de donner un sens aux informations. Elle est aussi un modèle de données qui représente un ensemble de concepts dans un domaine et les rapports entre ces concepts. Elle est employée pour raisonner au sujet des objets dans ce domaine.* » (Wikipedia). Les concepts sont ainsi organisés les uns par rapport aux autres via des relations de sémantique, des relations de composition et des relations d'héritage (cf. Figure 6).

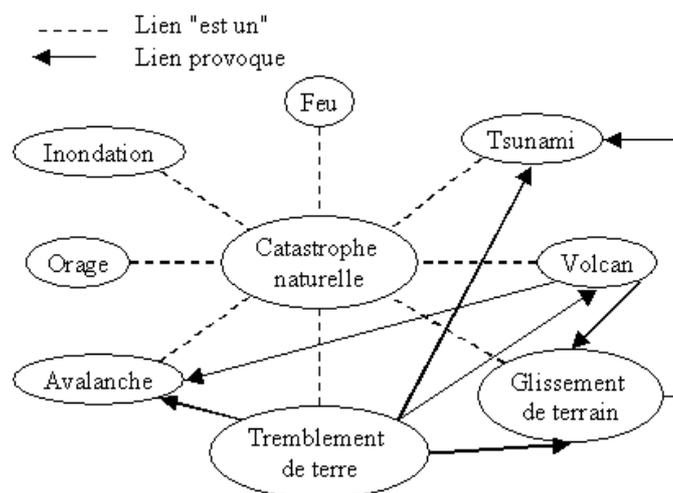


Figure 6 Représentation graphique d'une ontologie

Nous avons pu assister à un développement important des ontologies formelles et de technologies les supportant avec le mouvement lié au Web Sémantique (Berners-Lee et al. 2001). Ces ontologies permettent, par exemple, de structurer les concepts et la terminologie,

mais dans un domaine particulier. Elles sont majoritairement construites en s'appuyant sur des experts. Leur développement peut être rapproché de celui des systèmes experts. Elles sont d'ailleurs adaptées pour l'inférence et le raisonnement, permettant, par exemple, d'assister un moteur de recherche. En ce sens, elles ne représentent pas nécessairement des systèmes d'organisation de connaissances, se rapprochant parfois plutôt des objectifs des systèmes à base de connaissances dans leur mise en œuvre à l'issue de leur construction.

Une alternative aux ontologies que nous avons qualifiée de formelle correspond aux *ontologies sémiotiques* (Zacklad 2005). Ces ontologies s'inscrivent dans le développement du Web socio sémantique complémentaire du Web sémantique. Le Web sémantique « *met principalement l'accent sur la définition « d'ontologies formelles » (universelles et totalement standardisées) »* (Zacklad 2005b). Pour sa part le Web Socio Sémantique «  *vise à fournir des bases de codifications malléables et adaptées aux besoins évolutifs des communautés d'utilisateurs locales »* (Zacklad 2005b). Ce type d'ontologie doit permettre de mettre « *en relation interprétative à la fois les concepts sémiotiques et les situations qu'ils « dénotent » et les « connotations » mutuelles des concepts sémiotiques entre eux dans une logique multi points de vue »*. (Zacklad et al. 2003). Ainsi, la mise en relation des concepts sémiotiques avec les situations qu'ils dénotent est rendu possible par l'accès direct, via les concepts, à des documents qui les expliquent. Les connotations mutuelles s'expriment par des annotations portées par les multiples auteurs engagés dans le développement de l'ontologie sémiotique (Lortal et al. 2007; Zacklad 2007). Ce type d'ontologie résulte d'un processus d'indexation sociale, l'indexation étant une activité de repérage et d'identification d'information. Elle est sociale puisque, dans le contexte du Web socio sémantique, elle est pratiquée collectivement, sans être centrée sur une autorité (une institution ou un expert bien défini), ce qui la rapproche des folksonomies (Ertzscheid 2006).

Après avoir dressé un panorama général de l'activité de gestion des connaissances, de ses tenants et aboutissants, nous proposons d'en donner une vision plus concrète. La section suivante propose un aperçu de l'activité de gestion des connaissances en milieu industriel pour le cas particulier d'EDF.

### 3 - EDF et la gestion des connaissances en milieu industriel

Depuis maintenant une quinzaine d'année, entreprises et organisations ont largement pris conscience de l'enjeu crucial que représente la pérennisation des savoirs et savoirs faire qui représentent un capital immatériel de premier plan. La société EDF en tant que producteur d'énergie doit faire face à la gestion d'un vaste parc industriel. Ce parc se compose d'actifs de production et de transformation ainsi que d'infrastructures de transport de cette énergie. Au-delà de la dimension technique, c'est également tout un savoir-faire clairement multidisciplinaire qui est parsemé dans l'ensemble des personnels de l'entreprise et ce à différents niveaux hiérarchiques et dans différentes branches. Il s'agit donc à la fois de tenter de capitaliser les savoir-faire experts (Haïk et al. 2002; Haïk et al. 2003) notamment au regard de départs massifs en inactivité. Un enjeu majeur et incontournable est par exemple de « *préserver et exploiter les connaissances de diagnostic et de maintenance des composants exploités en centrale* » (Ricard et al. 2000) qui au fil du temps se sont dispersées dans les individus. Nous proposons de décrire plus précisément l'enjeu que représente la gestion des connaissances en milieu industriel.

#### 3.1. Un enjeu majeur

L'enjeu majeur que constitue gestion et ingénierie des connaissances pour EDF peut être abordé selon:

- Des Savoirs et Savoir-Faire techniques Répartis dans l'Espace et le Temps
- Des Savoirs et Savoir-Faire Volatiles
- Des Savoirs et Savoir-Faire Critiques pour la Sûreté
- Des Savoirs et Savoir-Faire Concurrentiels

Ces différents thèmes ne sont pas indépendants, certain partie prenante ou impactant sur les autres.

##### *3.1.a Des Savoirs et Savoir-Faire techniques Répartis dans l'Espace et le Temps*

Le domaine de la production d'énergie renvoie à un contexte éminemment technique. Les processus de production qu'ils soient liés à des moyens nucléaires, thermique à flamme, hydrauliques ou encore à l'éventail d'énergies renouvelables telles que le solaire ou l'éolien, mettent en jeu de nombreuses compétences différentes. Nous sommes alors confrontés à une activité profondément transverse mettant en jeu de nombreux acteurs. Les savoirs et savoir-

faire correspondent alors à un construit réparti dans les individus et les artefacts au sens de la cognition distribuée (Hutchins 1991; Smith 1994). Ainsi, chaque acteur possède une vision plus ou moins globale et plus ou moins précise d'un actif de production, du processus ainsi que de l'organisation dans laquelle il est plongé. Chacun introduit des niveaux et des types d'expertise différents (théorique, terrain) rendant l'identification et la maîtrise du processus de production parfois complexe.

Au-delà du caractère transverse et réparti des compétences mises en jeu s'ajoute une dispersion géographique. Le parc de production énergétique ainsi que les différentes directions (production, ingénierie, recherche et développement) d'EDF sont réparties sur l'ensemble du territoire. Ceci introduit une répartition dans l'espace des savoirs et savoir-faire de l'entreprise.

Enfin, il s'agit de ne pas négliger la dimension temporelle. En effet, les actifs de production ont une durée de vie particulièrement longue. Les moyens de production nucléaire actuellement mis en œuvre datent des années 70, la génération précédente (UNGG : Uranium naturel-graphite-gaz) ayant été exploitée durant une dizaine d'année en France. Les moyens hydrauliques (barrages, écluses, conduites forcées), quant à eux, peuvent dater de la fin du XIX<sup>ème</sup> siècle. La connaissance liée à ces ouvrages se construit donc dans la durée et sur l'ensemble de leur cycle de vie allant de la construction à l'exploitation en passant par le maintien et finalement le démantèlement.

On identifie bien ici le caractère réparti des savoirs et savoir-faire que cela soit sur la dimension humaine, spatiale et temporelle. Gestion et Ingénierie des connaissances doivent en ce sens apporter des éléments de solution à la fois techniques et organisationnelles permettant d'éviter une dilution de la maîtrise technique du parc de production. Le partage et la diffusion des informations capitalisées doivent également permettre d'améliorer la coopération entre les acteurs et les directions impliquées.

### *3.1.b Des Savoirs et Savoir-Faire Volatiles*

Nous l'avons évoqué la production d'énergie met en jeu de nombreux acteurs de domaines de compétences différents dans une activité relativement complexe. Chaque acteur constitue une source de compétence, de savoir, pouvant être critique pour l'activité. Ainsi, les changements d'acteurs représentent un source potentielle de déstabilisation. Ces changements peuvent aussi bien être liés à des réorganisations, des phénomènes de mobilité interne ou encore des départs

en inactivité. Ce dernier point se révèle critique dans les 5 années qui viennent étant donné le départ massif en relation avec le « baby boom » de l'après Seconde Guerre Mondiale qui s'est révélé particulièrement élevé en France.

La gestion des connaissances doit en ce sens favoriser l'échange en continu des savoirs et savoir faire. Ceci passe tout aussi bien par la constitution de base de connaissances qu'elle soit sous forme papier ou numérique, permettant de constituer une représentation pouvant être transmise. Il s'agit également de mettre en place des mécanismes sociaux permettant une transmission des savoirs, un exemple pouvant être le compagnonnage (Grassaud 2000).

### *3.1.c Des Savoirs et Savoir-Faire Critiques pour la Sûreté*

Le fonctionnement et l'évolution d'EDF sont très largement influencés par une problématique de sûreté. Les objectifs ne sont pas uniquement définis au regard de la rentabilité et du profit tel que cela peut être le cas dans bon nombre de secteur, mais doivent intégrer une réglementation gagée, par exemple, par l'autorité de sûreté nucléaire au niveau national et l'Agence Internationale de l'Energie Atomique (AIEA - <http://www.iaea.org/>), par exemple, au niveau international. Les objectifs de sûreté restent ainsi un enjeu prioritaire qui influence largement la stratégie d'investissement.

La gestion des connaissances joue un rôle majeur dans cette problématique puisque la maîtrise de la sûreté se fonde sur une connaissance complète du processus pouvant être très complexe ainsi qu'une connaissance des risques potentiels. Cette complexité, déjà évoquée précédemment, est liée à la diversité des acteurs et compétences mis en jeu, la répartition des savoirs dans le temps et l'espace, et également la complexité des équipements et de leur fonctionnement ainsi que des processus régissant la production énergétique. Dans ce contexte, le « besoin de connaissances » croît du fait de la variété et la complexité des savoirs et savoirs faire nécessaires (Ashby 1956).

### *3.1.d Des Savoirs et Savoir-Faire Concurrentiels*

Il est effectivement reconnu que savoirs et savoirs-faire représentent aujourd'hui un capital immatériel essentiel pour toute entreprise. La maîtrise de l'information représente un vecteur de développement et potentiellement une source d'innovation. Ceci explique notamment tout l'engouement et le développement des sociétés d'intelligence économique et de veille, mais également le courant du Knowledge management des années 90.

EDF est quant à elle, pour diverses raisons, a un tournant de son histoire. Tout d'abord, il s'agit de tenir compte du contexte énergétique mondial notamment au regard des problématiques environnementales et du réchauffement climatique. A cela s'ajoute la mise sur le marché d'une partie du capital d'EDF qui a eu lieu à la fin de l'année 2005 et qui transforme largement la physionomie de l'entreprise. Elle n'est plus uniquement ancrée dans une démarche de service public, mais entre clairement sur un marché concurrentiel. Enfin, l'ouverture à la concurrence du marché de l'électricité, suite aux directives européennes dans un premier temps aux professionnels, puis aux particuliers (juin 2007) influence nécessairement la stratégie économique.

Ce passage d'EDF de la sphère publique à la sphère privée remet en première ligne la problématique de la gestion des connaissances de l'entreprise. Les connaissances et informations détenues sont donc un vecteur potentiel de développement et d'innovation pouvant permettre d'obtenir un avantage concurrentiel important. Il s'agit donc d'en faciliter l'accès, l'utilisation voir la réutilisation pour supporter au mieux les activités de l'entreprise et favoriser l'atteinte d'objectifs stratégiques.

De plus, l'ouverture d'EDF induit également une logique de service. Les différentes branches de l'entreprise ne sont plus simplement centrées sur l'activité de production et de gestion de l'énergie. Des partenariats, des collaborations, peuvent ainsi être initiés avec d'autres secteurs d'activité. Il s'agit de s'écarter du principe de spécialité de la branche Recherche et Développement qui n'est plus uniquement centré sur les problématiques liées à la production et au transport du courant électrique. Il est alors essentiel de valoriser au mieux les savoir et savoir faire de l'entreprise afin de pouvoir valoriser ses compétences. Ceci implique d'être capable de les recenser et donc de les identifier. On assiste déjà à ce genre d'ouverture au niveau recherche et développement au travers de ce que l'on nomme la « R&D dédiée ». La branche Recherche et Développement s'insère dans une logique de service (Zacklad 2007) pouvant être valorisée sous la forme de prestations au-delà du cercle de l'entreprise. Ceci peut être fait avec d'autres acteurs internationaux de l'énergie (Westinghouse pour les Etats Unis, Tepco pour le Japon ...) au travers d'un partage d'expériences, d'outils ou de méthodologies. Il peut également s'agir de fournir un savoir faire, une expertise, à des entreprises d'autres secteurs d'activité et ce dans une logique de rentabilisation et d'ouverture de la branche R&D.

La section suivante propose d'illustrer l'enjeu industriel à travers un panorama de l'activité de gestion des connaissances telle qu'elle est pratiquée au sein d'EDF.

## 3.2. Quelques exemples

Les enjeux sous jacent à l'activité de gestion des connaissances dans un groupe comme EDF fait que les exemple sont nombreux. Nous pourrions citer la construction d'ontologies relatives au domaine de l'hydraulique (Dourgnon-Hanoune et al. 2006), ou encore une problématique de mise en relation de données physiques relatives à des mécanismes de dégradations avec des documents de capitalisation se matérialisant sous la forme d'un hypertexte (Mahé 2004; Haïk 2005). Nous proposons ici de décrire 2 exemples :

- La capitalisation des pratiques de maintenance dans le domaine de l'hydraulique : le développement de guides
- Méthodes et Outils d'Aide à la Gestion des Actifs de Production

Ces deux exemples offrent un bon aperçu de la problématique de l'inscription d'une représentation des connaissances sur un support, que cela soit pour une manipulation par un utilisateur ou une valorisation par la machine. Ils représentent notre terrain d'expérimentation. Chaque exemple est introduit par la problématique sur laquelle il se focalise et le contexte dans lequel elle s'inscrit. Les artefacts produits ainsi que les outils exploités sont également décrits.

### *3.2.a La capitalisation des pratiques de maintenance dans le domaine de l'hydraulique : le développement de guides*

Dans le cadre du développement des guides de diagnostic, la demande était guidée par le besoin de mettre à disposition du parc hydraulique des produits et services d'aide à la maintenance et d'aide à la décision. En prolongement de cette demande initiale, il s'agissait de développer des méthodes et outils pour la gestion des risques liés au patrimoine hydraulique.

Les ouvrages de production « hydrauliques » se caractérisent par une durée de vie particulièrement longue. Barrages, écluses ou encore conduites forcées voient leur longévité s'étaler sur plusieurs décennies. Il s'agit d'un cas typique dans lequel on identifie assez bien les enjeux évoqués dans la section 3.1. , page 39. En effet, tenant compte de la durée de vie des ouvrages ainsi que de leur dispersion géographique, il existe effectivement une répartition des savoirs et savoirs faire dans le temps et l'espace. Cette répartition se fait également ressentir au niveau des personnels étant donnée la durée de vie, mettant en avant le caractère très volatile des connaissances métier. Les enjeux de sûreté sont en relation avec une

réglementation à la fois imposée par les « Voies navigables de France » (VNF - <http://www.vnf.fr/>) et par la « Directions Régionales de l'Industrie, de la Recherche et de l'Environnement » (DRIRE - <http://www.drire.gouv.fr/>), mais également alimentée par la politique de l'entreprise. Enfin, l'enjeu concurrentiel est également présent en relation avec l'augmentation du transport fluvial (nombre de navires) que cela soit en fréquence ou encore en volume (taille des navires).

L'enjeu patrimonial du projet est lié au vieillissement du parc hydraulique. De nombreuses techniques de construction cohabitent, renvoyant à des conceptions différentes d'ouvrages. Cette diversité des connaissances mises en jeu est renforcée par le fait qu'en général au moins 3 domaines de compétences cohabitent dans l'exploitation des ouvrages : le génie civil, l'hydro mécanique et le contrôle commande. Face à cette diversité, les opérationnels engagés dans l'exploitation et le maintien des ouvrages ont développé des techniques et stratégies de maintenance qui sont parfois propres à certains ouvrages, ce qui peut être justifié du fait d'une spécificité, mais qui pourrait également mériter d'être partagé et diffusé. La diversité des technologies déployées complique également l'analyse du vieillissement des structures de génie civil et mécanique.

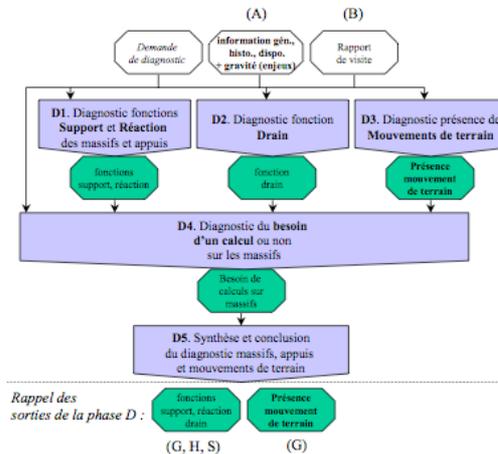
Au travers de ce projet, le travail s'est porté vers la construction d'un référentiel commun développé en s'appuyant sur les acteurs impliqués dans la maintenance de l'ouvrage. Au travers de la construction de ce référentiel, c'est également la notion de diagnostic qui a été mieux appréhendée et définie. Le diagnostic est ainsi une représentation qui tend à mobiliser à la fois des compétences de terrain, liées à l'observation de l'ouvrage par exemple, ainsi que ses connaissances théoriques de l'ouvrage. Le développement du référentiel avait donc pour objet d'articuler ces différentes informations.

Dans le cadre du développement des guides, la ressource produite correspond à un document dont la forme et les outils exploités pour le produire ont évolués. La construction du guide s'est donc faite au travers d'une succession d'outils bureautiques passant du traitement de texte Microsoft Word (cf. Figure 7) au logiciel de construction de diagramme Visio (cf. Figure 8), suivi d'un déploiement au travers du format PDF avant d'intégrer une logique plus structurée basée sur l'exploitation d'éditeurs XML (Parfouru et al. 2006; Parfouru et al. 2006)

- fonction support (utile pour G et S)
- fonction réaction (utile pour G et S) : blocage
- fonction drain (utile pour G et S)
- présence de mouvement de terrain (utile pour G et S)
- besoin de calcul de massifs et appuis (utile pour H)

NB. les calculs des massifs et appuis sont réalisés dans la phase H qui est optionnelle (cf. 2.7 « Phase H. Calcul de massifs et appuis » page 36).

### 2.4.1 Synthétique de la phase D



Le synoptique de la phase D est élaboré pour les conduites aériennes car elles représentent la majorité de notre parc. En ce qui concerne le **cas particulier des conduites enterrées**, l'expert devra orienter son diagnostic de cette phase avec la même démarche intellectuelle, mais en prenant en compte les spécificités particulières des CF enterrées (notamment les efforts dus aux charges de remblais et divers).

Figure 7 – Le guide de diagnostic des conduites forcées construit au travers du traitement de texte Microsoft Word

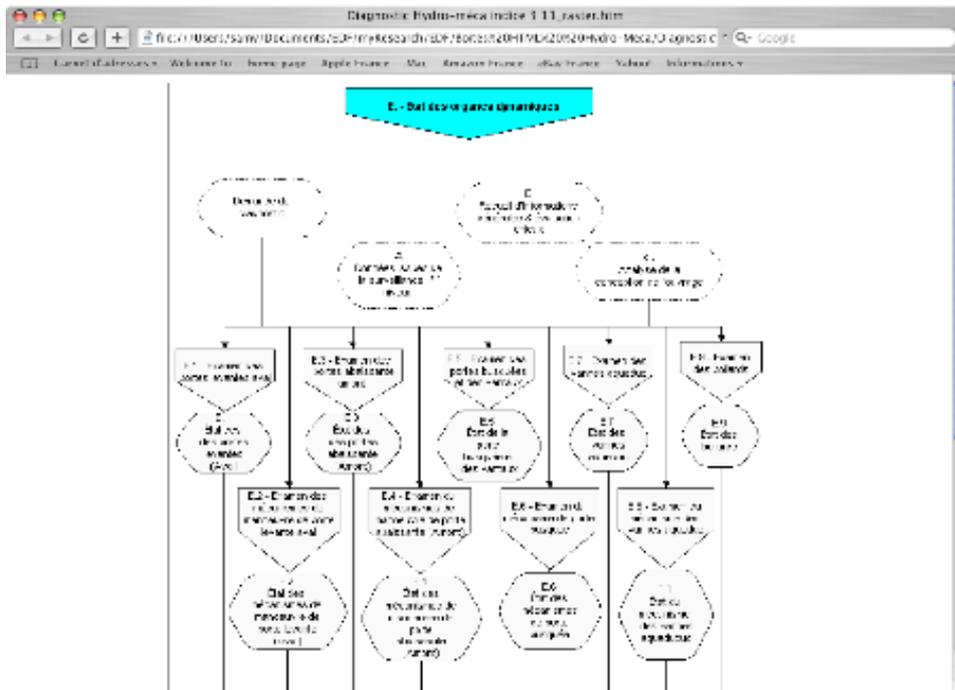


Figure 8 – Le guide d'aide au diagnostic des écluses construit via microsoft Visio et déployé via navigateur WEB

### *3.2.b Méthodes et Outils d'Aide à la Gestion des Actifs de Production*

La longévité des actifs de production en présence pose une problématique particulièrement complexe aux décideurs. Il s'agit de considérer de nombreux éléments liés à la prise en compte d'aspects techniques, économiques et politiques qui prennent un caractère très incertain sur des périodes d'exploitation aussi longues. Il est en effet difficile à moyen ou long terme de prévoir, par exemple, les contraintes de sûreté imposées par une autorité indépendante, les évolutions technologiques ayant un impact sur la productivité du parc ou encore les changements de stratégie de l'entreprise. La définition des stratégies d'exploitation et de maintenance qui prennent un caractère crucial, dans un environnement concurrentiel, sont particulièrement difficiles à déterminer sachant qu'il s'agit à la fois d'optimiser la durée d'exploitation de l'actif tout en répondant, en priorité, aux contraintes de sûreté.

En ce sens, il s'agit de développer des méthodes et outils permettant d'appuyer l'activité des décideurs et la définition de ce type de stratégie. Ce type d'aide s'appuie à la fois sur une bonne connaissance des ouvrages, des technologies mises en œuvre ou encore des mécanismes de vieillissement actuellement connus. Il s'agit également de prendre en compte le retour d'expériences relatif au parc : celui-ci à la fois renvoie aux problèmes déjà rencontrés, mais également aux politiques de gestion actuellement mises en œuvre. Enfin, il s'agit de pousser les experts à déterminer les risques potentiels, notamment au travers de méthodes d'aide à la créativité (Bouzaiène et al. 2002; Haïk 2004; Bouzaiène 2005) comme dans un processus d'innovation, et ainsi déterminer les actions de maintenance permettant de s'en prémunir.

L'étude engagée dans cette problématique a permis de développer un modèle probabiliste permettant à partir d'une représentation formalisée de connaissances de calculer des scénarios possibles de l'évolution des actifs de production (Bauby et al. 2004; Bauby et al. 2005). Cette approche basée sur la simulation permet d'offrir une vision calculée et probabilisée de la valeur d'un actif de production (pour un contexte d'exploitation et de maintenance) afin de faciliter la prise de décision et le développement de politique de maintenance à long terme. Elle ne se concentre toutefois pas uniquement sur le développement d'une représentation des connaissances manipulables par la machine.

En effet, une méthode d'instruction a été développée qui tente d'associer à la fois une approche basée sur la rédaction de documents de capitalisation construits sur la base d'entretiens d'experts et de données de retour d'expérience (repère 1 sur la Figure 9). Ces

documents écrits, qui représentent une forme de représentation de l'information appréhendable par l'ensemble des acteurs – contrairement à une représentation formelle adaptée à un calcul algorithmique –, prennent alors un caractère de document de référence sur l'objet, la problématique, sur laquelle ils se concentrent.

En prolongement de ces documents de référence, une modélisation formelle de leur contenu est produite (repère 2 sur la Figure 9), correspondant à la fois à un échantillonnage de leur contenu, ainsi qu'à une reformulation, une réécriture, dans un langage adapté à un système informatique de simulation (réécriture réalisée par un analyste ayant une expertise sur le système). Ce système permet, à partir de techniques statistiques et probabilistes, de calculer l'évolution probable d'indicateurs technico-économiques (repère 4 sur la Figure 9) ainsi que les scénarios fournissant une image des possibles évolutions de l'actif de production analysé.

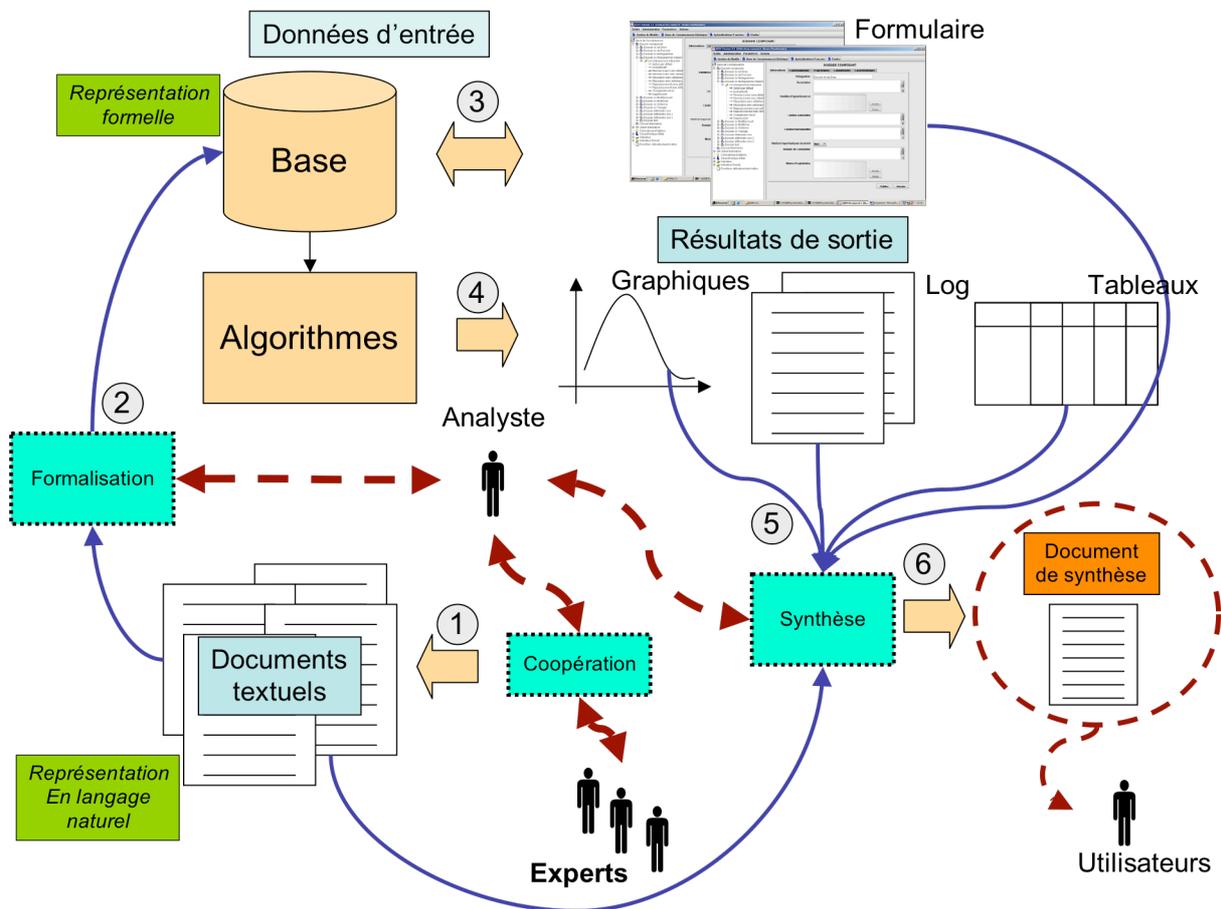


Figure 9 Vue schématique du contexte organisationnel et technique du projet

Ce projet est clairement multi disciplinaire. Ainsi, le développement de la méthode et de l'outil font appel à la fois à des compétences en statistiques et probabilités (pour la définition des mécanismes de simulation), à l'ingénierie des connaissances (pour la capitalisation

centrée alors sur le recueil et la modélisation). Ces compétences sont également essentielles face à l'interprétation des résultats produits. Il s'agit de s'intéresser à des connaissances détenues à la fois par des experts techniques des actifs de production analysés, par des financiers (pour appréhender la logique économique et d'investissement), et par les décideurs (pour leur vision stratégique).

Le projet a donné lieu au développement d'une plate-forme informatique permettant à la fois de saisir les connaissances sous une forme manipulable par la machine tout en intégrant l'algorithmique de simulation (repère 3 sur la Figure 9). Ceci étant, de nombreux documents, produits via des outils d'informatique bureautique, gravitent autour du système et sont intégrés à la méthodologie d'instruction. Il s'agit, par exemple, des notes de capitalisation de connaissances, rédigées en collaboration avec les experts et qui prennent un réel statut de documents de référence au regard de leur contenu et du sujet qu'elles traitent. La complexité inhérente aux résultats produits (repères 5 et 6 sur la Figure 9) par le système oblige également à produire des documents à l'issue des simulations. Ces notes de synthèse des résultats mettent ainsi en relation, par exemple, les connaissances techniques, financières et stratégiques capitalisées, l'évolution des indicateurs calculés et les éléments constituant la simulation afin de produire un résultat argumenté.

## Chapitre 2 - Problématique

### 1 - L'objet d'étude

L'étude proposée dans ce mémoire s'intéresse à l'activité de gestion des connaissances dans un milieu industriel. Il s'agit d'analyser les pratiques de capitalisation et de restitution des connaissances d'une équipe dont la tâche est de préserver une partie des savoirs et savoir faire métier d'EDF. Cette équipe se compose de personnels d'origines et de domaines de compétences très différents comme la sociologie, les sciences de gestion ou encore l'informatique. Cette diversité de compétences, qui pourrait laisser entrevoir un manque de cœur de métier ou de domaine de référence dans le cadre de l'activité de gestion des connaissances, est en fait essentielle. Dans ce type de démarche qui est, par nature, transdisciplinaire, la diversité que nous avons évoquée permet de multiplier les points de vue et les approches, ce qui permet de créer le débat pouvant générer une dynamique féconde, en particulier, dans l'approche méthodologique.

L'approche méthodologique développée au sein de la compétence ne peut être dissociée du contexte socio-organisationnel dans laquelle elle s'insère. La gestion des connaissances telle qu'elle est pratiquée au sein d'EDF va donc mobiliser un réseau d'acteurs plus ou moins étendu, mais que l'on peut schématiser sous la forme de la Figure 10.

Il se compose :

- Des Analystes,
- Des Opérationnels,
- Des Commanditaires,
- Des Informaticiens.

Comme l'illustre la figure, le réseau d'acteur est centré sur l'*analyste*. Celui-ci mène la démarche de gestion des connaissances et représente en ce sens un élément pivot de la démarche. Les analyste correspondent ici à des personnels de la branche Recherche et Développement même s'il est tout à fait admis que de nombreux personnels pratiquent une forme de gestion des connaissances sans nécessairement le savoir (Dieng-Kuntz et al. 2001). L'analyste est la personne qui va mener la démarche de capitalisation des connaissances. Il

doit cultiver la dynamique de cette démarche et, nous le verrons par la suite, y joue un rôle de médiateur. Il est en charge de :

- L'identification des experts.
- La co-construction des modèles en collaboration avec les experts.
- La conduite des réunions de capitalisation afin d'accompagner les experts dans la formalisation des connaissances en fonction des possibilités des modèles.
- Dans le cadre du développement d'un système informatique, il pourra éventuellement s'attacher à la spécification des algorithmes.

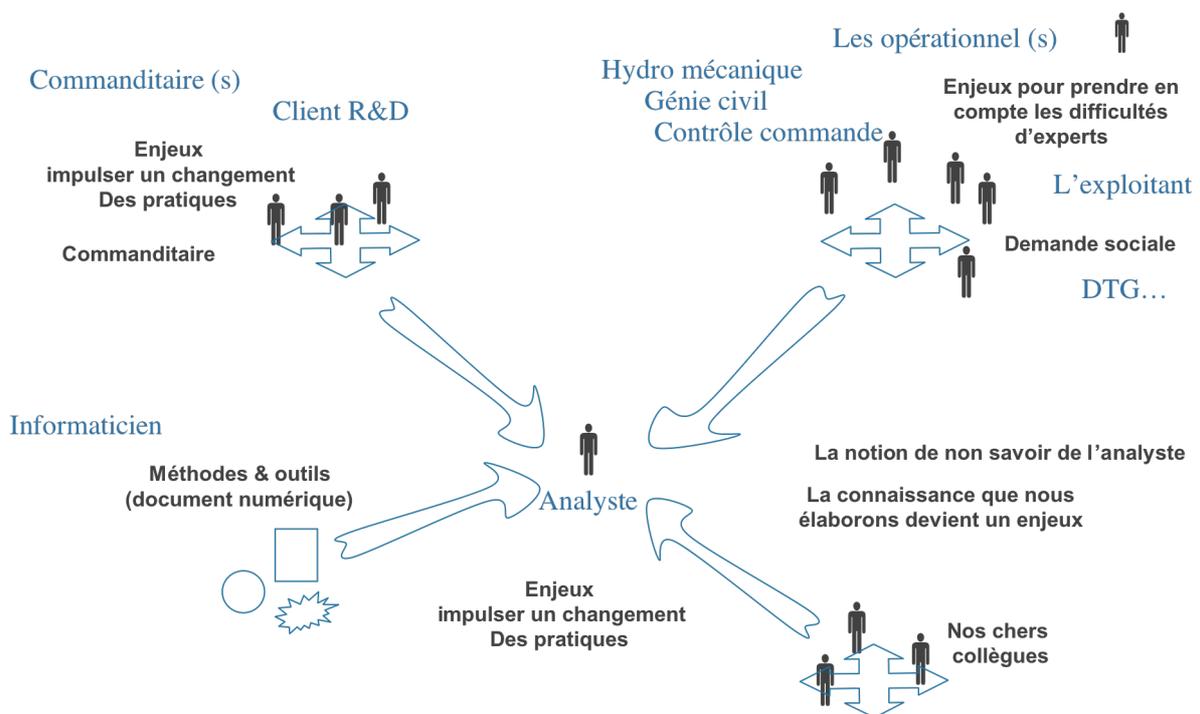


Figure 10 - Réseau d'acteurs dans lequel s'insère une activité de gestion des connaissances

L'activité de gestion des connaissances n'est pas inscrite dans une démarche globale et transverse à l'ensemble de l'entreprise, elle s'insère donc, le plus souvent, dans une logique de projets en référence à des problématiques liées à l'entreprise. Ces problématiques sont soumises aux analystes par des commanditaires. Ces *commanditaires* au regard du projet introduisent une vision globale et stratégique en relation avec le développement de l'entreprise ou de l'activité. Les enjeux renvoient aux changements de pratiques, à la performance, etc... Le commanditaire est donc porteur de la problématique à traiter. Il peut être à l'origine de cette problématique, mais pas toujours. C'est un acteur tout à fait important

avec qui il convient de savoir négocier. En effet, le commanditaire peut avoir une vision à priori en relation avec ce à quoi il a déjà été confronté, ce qu'il est capable d'imaginer, qui peut se révéler en contradiction avec une réalité de terrain identifiée par l'analyste.

Une catégorie d'acteurs incontournable est celle des *opérationnels* dans laquelle nous intégrerons ceux communément appelés *experts*. La notion d'expert reste une notion floue, l'expertise pouvant prendre de nombreux visages. Elle peut être liée à une connaissance théorique approfondie ou à une longue expérience de terrain. Ce statut d'expert véhicule une reconnaissance dans la communauté et une légitimité au regard de l'information fournie. Dans le cadre d'un vaste projet de gestion des connaissances, on tentera classiquement de faire appel à plusieurs experts, chacun apportant un point de vue différent sur le ou les domaines que l'on cherche à formaliser. Les opérationnels renvoient également aux futurs utilisateurs du produit de la démarche engagée. Que la finalité soit la production d'un document ou bien un système d'information, un projet d'Ingénierie des Connaissances fait référence à un ou plusieurs opérationnels qui exploiteront le résultat du projet dans des pratiques métiers.

Les projets de gestion de connaissances s'appuient nécessairement aujourd'hui sur les technologies de l'information et de la communication. Même dans la rédaction d'un document, ce dernier prend nécessairement aujourd'hui une forme numérique à un moment de son cycle de vie. S'il est toujours possible pour certains analystes de conserver une autonomie dans le développement des ressources numériques produites, qu'ils s'agissent de documents numériques ou de bases de données, beaucoup de problèmes peuvent se poser si la ressource en question prend de l'ampleur. Il est alors incontournable d'intégrer les *informaticiens* dans le processus de gestion des connaissances. Ces derniers apportent un appui, tout d'abord sur la dimension technique et en particulier les technologies disponibles et adaptées à la problématique posée. Ils introduisent également un savoir faire et une culture de la gestion, du maintien et de la préservation d'informations numériques.

La logique de projet dans laquelle s'insère l'activité de gestion des connaissances peut introduire des utilisateurs potentiels ou finaux du produit résultant de la démarche permettant de définir des besoins. Ceci étant, cet acteur peut être parfois difficile à identifier. La gestion des connaissances correspond à une activité peu déterministe au sens où identifier la ressource produite avant l'issue du processus peut se révéler complexe. De ce fait, il est difficile de réellement pouvoir identifier l'ensemble des utilisateurs en amont. De plus, l'observation montre que l'ensemble des acteurs inscrits dans ce réseau devient usager de la ressource

produite à un moment ou à un autre que cela soit pour la concevoir, la développer, l'alimenter, la valider ou encore l'exploiter. Les utilisateurs sont alors nombreux, ce qui peut largement compliquer la définition des besoins en particulier sur les moyens d'accès à l'information et sur la forme qu'elle doit prendre.

## 2 - Les problématiques

Au regard de l'objet d'étude que nous venons de décrire, nous nous intéressons à deux aspects :

- La capitalisation des connaissances
- La restitution d'informations

Il s'agit de deux problématiques très vastes. Aussi, l'ambition n'est pas de vouloir les traiter dans leur intégralité.

### 2.1. L'activité de capitalisation de connaissances

L'étape de capitalisation des connaissances est décisive dans les démarches de gestion des connaissances. Elle tient à beaucoup d'éléments. Il s'agit tout d'abord d'identifier les acteurs devant être impliqués (même si les participants peuvent être éventuellement modulés au cours et au regard de l'évolution du projet). L'ensemble des participants va constituer un réseau d'acteurs engagés dans la réalisation d'un projet commun relatif à la définition d'une ressource pouvant tout aussi bien faire référence à un document ou un système informatique intégrant et véhiculant une représentation des connaissances dont les modalités sont à définir.

Le bon déroulement de cette activité dépend du tissage d'un réseau social entre de nombreux acteurs. Ces différents acteurs peuvent être en accord ou non. Le bon déroulement de la démarche de capitalisation tient ainsi au fait qu'elle doit mettre en relation ces acteurs pour qu'ils adhèrent à un projet commun. Notons bien que ce projet commun est identifié au cours de la capitalisation elle-même, et doit faire le lien entre les besoins exprimés par le commanditaire et ceux qui émergent du terrain via les interactions avec les opérationnels. Ainsi, il ne s'agit pas de faire uniquement appel aux acteurs en les considérant comme fournisseurs d'informations, mais de totalement les intégrer à la conception du produit vers lequel on tend. En ce sens, les démarches de gestion des connaissances peuvent parfois même réconcilier certains acteurs et faire qu'ils coopèrent mieux à l'avenir.

Le tissage d'un réseau social entre les acteurs est en relation avec leur adhésion à la démarche de gestion des connaissances. De là dépend la qualité de la ressource, de l'artefact, que l'on cherche à produire et notamment la qualité et la précision des données. En arrière plan de la capitalisation des connaissances, c'est un enjeu de pouvoir qui se joue. La capitalisation tient à la fois à la volonté de partage des acteurs mais également de la contrepartie qu'ils pourront avoir s'ils fournissent des informations. Il s'agit donc que la capitalisation aboutisse à un produit qui puisse s'intégrer dans leurs activités et qui puisse, également, être valorisé auprès d'autres acteurs.

Une grande part de travail de l'analyste est donc de valoriser au mieux le réseau d'acteurs dans lequel il est lui-même plongé. Ce réseau introduit à la fois un contexte multidisciplinaire et multi point de vue. Ce contexte, également en relation avec le caractère intrinsèquement transverse de la problématique de gestion des connaissances, amène plusieurs difficultés dans :

- La définition de la représentation de connaissances
- La définition du produit de la démarche de gestion des connaissances
- L'impact de la démarche sur les acteurs

### *2.1.a Définition de la représentation de connaissances*

Dans la définition de la représentation des connaissances, chaque domaine va introduire des langages différents au regard de la terminologie employée ou encore des modes de représentations exploités. Ces modes de représentations qui introduisent de réelles codifications, certains domaines s'appuyant sur des représentations schématiques, exploitent une symbolique particulière et peuvent être le résultat de nombreuses années d'existence du domaine. Ainsi, en électrotechnique certains symboles qui originellement dans leur visualisation imitaient le fonctionnement du mécanisme ont par exemple été conservés malgré l'évolution technologique du fonctionnement de ce mécanisme.

Au-delà des modes de représentation de l'information, chaque domaine de compétence véhicule des habitudes, des conventions, un regard, ou point de vue, sur l'objet auquel on s'intéresse. Ainsi, une personne issue du génie civil pourra appréhender l'ouvrage selon son architecture, alors qu'en hydro-mécanique la vision sera influencée par la chaîne cinématique de l'ouvrage, ou encore en contrôle commande on s'intéressera majoritairement à l'ouvrage selon sa dynamique de commande. Bien évidemment certaines parties d'un ouvrage sont

propres à un domaine, mais il existe toujours des points de contact entre des éléments nécessitant le regard de différentes compétences. Il convient donc dans la démarche de capitalisation de tenir compte de ces aspects multi domaine et multi point de vue car la participation et l'implication des acteurs dépend largement du degré de reconnaissance qu'il auront dans la représentation des connaissances produites. Ainsi, chaque domaine doit pouvoir se reconnaître, mais également appréhender les recoupements avec les autres compétences.

### *2.1.b Définition du produit de la démarche de gestion des connaissances*

L'aspect multi point de vue doit également être pris en compte au regard de la finalité du projet de gestion des connaissances. Qu'il s'agisse de produire un document ou un système informatique, la vision des acteurs impliqués peut être très différente sur ses objectifs et ses usages. Ainsi, si un décideur pourra vouloir développer un système plutôt prescriptif, les personnels de terrain pourront être très réfractaires à cette idée, étant à la recherche d'un support à leur activité comme par exemple un aide mémoire (Leplat 2004). De même, au delà de l'objectif de la ressource, les différences de point de vue peuvent se ressentir au niveau des modalités de sa définition : quelle support (medium) ? quelle forme ? quel niveau de précision ? quel niveau de structuration ?

A cela s'ajoute le fait que les membres du réseau d'acteurs mobilisés ne sont pas toujours au fait de ce qu'il est effectivement possible de réaliser. C'est notamment le cas dans la réalisation de ressources ou systèmes informatiques. La démocratisation de la micro-informatique ainsi que l'accès à Internet a permis aux futurs utilisateurs de mieux appréhender les possibilités de l'ordinateur, ceux-ci devenant, par exemple, exigeants face aux interfaces proposées (Gehland et al. 2006). Cependant, ils ne possèdent pas toujours une vue claire sur les réalisations possibles dans le cadre de traitements complexes. Ils ont parfois une vue faussée face au développement d'application ou de fonctionnalités dont ils n'appréhendent pas la complexité. Ainsi, il peut être difficile, en particulier dans le développement d'un système informatique, de définir le produit de la démarche de gestion des connaissances a priori nous plaçant dans une activité d'innovation. Ainsi, « *le projet, comme dans toute innovation, se construit en même temps que se fait le choix technique.* » (Boullier 2001).

### 2.1.c Impact de la capitalisation sur les acteurs

L'activité même de capitalisation des connaissances a un impact fort sur les acteurs qui y sont engagés puisqu'elle influence nécessairement leurs propres connaissances. En effet, il s'agit de « *comprendre les changements survenus au cours des projets en insistant sur le processus de métamorphose des acteurs qui est corrélatif à la recomposition des savoirs opérés* ». (Hatchuel et Weil 1992). La recomposition des savoirs est liée au fait que l'identification et la formalisation de connaissances va réinterroger chacun des participants sur son propre savoir. Ainsi, les différentes visions des acteurs liées au contexte multi domaine et multi point de vue au regard de la représentation des connaissances ou du produit de la démarche de gestion des connaissances sont changeantes tout au long du projet car influencées par les résultats intermédiaires présentés aux acteurs.

Ces résultats peuvent tout aussi bien faire référence aux informations capitalisées, représentant le plus souvent un savoir tacite qui réinterroge chaque acteur sur sa propre pratique, ou aux moyens, technologiques ou non, exploités pour concevoir et développer un artefact, leur permettant de mieux identifier ce qui est effectivement réalisable. L'ensemble des éléments présentés lorsqu'ils vont être assimilés par les acteurs constituent de nouvelles connaissances. « *La notion de métamorphose des acteurs est donc intimement liée à l'émergence de nouvelles connaissances et à leur distribution parmi les individus concernés.* » (Hatchuel et Weil 1992). Notons que « *cette dynamique n'est pas sans conséquence sur les réseaux d'acteurs concernés par le projet ; avec la restructuration des savoirs s'amorcent des tensions qui assureront sa fécondité ou son enlisement. Nous sommes alors au cœur de ce qui se joue dans un collectif lorsqu'il s'agit de créer ou de modifier la répartition des savoirs de chacun et donc le contenu des savoirs partagés.* » (Hatchuel et Weil 1992).

### 2.2. Restitution d'informations

Si la section précédente a mis en évidence certaines difficultés relatives à la capitalisation des connaissances, la restitution de la représentation produite n'est guère une tâche plus facile. Une définition de la restitution pourrait être « *la transformation de l'information interne sous une forme acceptable par les canaux de sortie* » (Nigay et Coutaz 1993). Dans cette définition, issue du champ disciplinaire de l'interface homme machine, l'*information interne* correspond à une information enregistrée dans un ordinateur, un support numérique, et les *canaux de sortie* font tout aussi bien référence à l'interface qu'elle soit d'interconnexion avec

un autre matériel ou de transmission avec un humain. La diffusion de l'information entre deux éléments est appréhendée comme une simple transformation de l'information introduisant des opérations de codage et décodage dans l'esprit de la théorie de l'information (Shannon et Weaver 1947). Cette vue est assez peu en adéquation avec notre contexte d'étude puisque les informations véhiculées ne peuvent être uniquement prises en compte au niveau signal, ce qui limiterait alors la « restitution » à une transformation de l'information pour une mise en adéquation avec les propriétés d'un support de communication.

Dans notre étude, la problématique de restitution est moins déterministe du fait que le problème et de restituer une information à un utilisateur humain. Celui-ci ne peut être appréhendé comme un objet modélisable et déterministe puisque face à une information, il initiera nécessairement un processus d'interprétation avec ses propriétés personnelles. Il n'y a alors pas de garantie qu'il appréhende correctement (c'est-à-dire tel qu'attendu par le concepteur du système de restitution) le sens de l'information qui lui est proposée.

Ainsi, dans un contexte de gestion des connaissances « *la diffusion d'une information ne suffit pas à garantir la réutilisation de la connaissance qu'elle est susceptible de transmettre. En effet, pour qu'une connaissance soit réutilisée, il est nécessaire qu'elle soit assimilée par l'acteur c'est à dire intégrée à sa base de connaissances et mobilisée à tout moment dans l'action* » (Castillo et al. 2004). Dans la même lignée, « *Il ne suffit pas que des connaissances soient disponibles pour être assimilées par les humains. Il ne suffit pas d'acheter des encyclopédies pour que les gens apprennent ! La connaissance disséminée sur les réseaux ne peut être réappropriée par les humains que si un effort est fait du point de vue didactique, dans la conception de situations d'apprentissage et de préceptorat* » (Castillo et al. 2004). Ainsi, il ne suffit pas de mettre à disposition ou de diffuser une information pour qu'elle puisse être exploitée et valorisée dans une activité. Il s'agit alors d'accompagner l'*appropriation*.

« *L'appropriation consiste dans le fait de faire sien le contenu et de l'intégrer comme une part de soi : penser ou agir avec lui, sans avoir à s'y référer explicitement. L'appropriation connaît deux modalités essentielles : l'appropriation comme application du contenu, c'est-à-dire une mise en œuvre ou une exécution, comme un mode d'emploi indiquant les actions à suivre ; l'appropriation comme reformulation ou réinvention de son expression, comme dans le commentaire ou la compréhension du contenu. Le contenu est donc approprié par le geste (application) et par la mémoire (reformulation).* (Bachimont et Crozat 2004). Ce passage met en avant 2 modalités relatives à l'appropriation que sont l'*application* et la *reformulation*. Ces deux modalités sont totalement en relation avec l'acquisition des connaissances, la

reformulation pouvant faire référence à l'acquisition d'un savoir, et l'application, dans une dimension plus pratique à l'acquisition d'un savoir faire.

De là, l'appropriation représente un élément central dans les démarches de gestion des connaissances en particulier dans la diffusion des informations capitalisées. Il s'agit donc de bien prendre en compte cet aspect dans la phase de restitution. Nous proposons de mettre en avant les difficultés relatives à l'appropriation de la représentation des connaissances. Nous avons vu précédemment que la définition de la représentation des connaissances était une tâche complexe influencée par le contexte multi domaine et multi point de vue dans lesquels elle s'insère. Définir la représentation des connaissances renvoie donc à une problématique de définition de langage devant prendre en considération les conventions et formalismes existants, mais également parfois en concevant des codes inédits, supports à la création d'un objet commun pour aboutir à un consensus entre compétences. Au regard de l'appropriation de la représentation des connaissances, cette dernière pouvant prendre la forme d'un document ou d'une modélisation formelle, la difficulté est que l'utilisateur de cette représentation soit en mesure de correctement l'interpréter.

Pour accompagner cette interprétation, il s'agit à la fois de bien faire transiter le langage de représentation défini et utilisé. Il s'agit également de prendre en considération les utilisateurs cibles de cette représentation qui peuvent être issues de domaines différents tout en introduisant des niveaux d'expertises variables. La problématique est alors de travailler la présentation des informations que cela soit au niveau de la forme et/ou du fond.

Dans (Castillo et al. 2004), la problématique de l'appropriation des connaissances est abordée en s'intéressant aux mémoires métier. On évoque que « *la présentation des mémoires-métier actuellement ne permet pas un accès facile à la connaissance* ». La difficulté peut s'expliquer comme tenant à deux aspects (Castillo et al. 2004) :

- « *les connaissances sont présentées sous plusieurs angles de vues (classifications, contraintes, processus, stratégies de résolutions de problèmes,...) = forment un frein à l'apprentissage* »
- « *la présentation qui en est faite est souvent en rapport avec une saisie procédurale ou une présentation brute des structures mises en jeu* »

Le premier aspect met l'accent sur la multiplicité des formes de représentation de connaissances dans le cadre des mémoires de projet. Cette multiplicité, largement favorisée

par la numérisation du support, constitue à la fois une force permettant de véhiculer différents points de vue. Elle est également une limite au sens où l'utilisateur qui va naviguer dans différents modes de représentation, passer d'un mode à un autre, va pouvoir développer une désorientation dans l'espace informationnel qui lui est proposé. Le fait de présenter par exemple les différents angles de vue de manière indépendante sans proposer une ligne directrice, une vision intégrée ou encore un accompagnement à la lecture peut alors constituer un « *frein à l'apprentissage* ». Il s'agit donc lorsque l'on restitue la ou les représentations de connaissances de produire un message adapté et ne pas se contenter de mettre à disposition une masse d'informations. Il est possible ici de faire le parallèle avec l'ingénierie pédagogique où l'activité de production de messages est présentée comme : « *Appareiller l'activité constituée par la production de messages, ce n'est plus gérer un stock de cours que l'on met en ligne, c'est assister le travail distribué de déconstruction / reconstruction des savoirs qui constitue l'enseignement.* » (Boullier 2001). Au travers de cette citation apparaît l'importance de produire une restitution propice à la manipulation par l'utilisateur. Il ne s'agit pas uniquement de diffuser un message figé, mais bien de produire une restitution dans laquelle on puisse naviguer, découvrir, et qui puisse s'intégrer dans l'activité de l'utilisateur : le message ne doit pas constituer un prescrit mais bien une « *assistance* ».

Le second aspect, lié aux systèmes informatiques, tient au fait que souvent la restitution (la visualisation qui est faite des informations) est calquée sur les structures de données implantées (« *présentation brute des structures mises en jeu* »). Un exemple caractéristique est le formulaire qui est une forme d'interface classique (cf. Figure 11). Les formulaires permettent et ont l'avantage de cadrer la saisie des utilisateurs afin qu'ils se conforment aux structures de données mises en jeu dans le système (Bachimont 2004). Cet intérêt des formulaires peut constituer une difficulté puisque les structures de données sont développées avec un souci de répondre aux contraintes de repérage, de cohérence et d'unicité des données. De ces contraintes résultent un découpage et des regroupements d'informations aisément compréhensibles lorsque l'on connaît l'implantation en machine, mais qui ne sont pas toujours propices à l'activité dans laquelle le système doit s'insérer.

La compréhension des structures de données mises en jeu constitue donc le premier problème. Ces structures sont en relation avec les modèles conceptuels développés, sur lesquels elles se fondent. De ce fait, il est indispensable d'avoir une connaissance de ces derniers pour les comprendre. Ceci nous amène à aborder la problématique des systèmes à base de connaissances et en particulier lorsqu'ils produisent des résultats basés sur un algorithme.

La compréhension des structures de données ne constitue pas le seul problème lorsque l'on est confronté à un système informatique. Les systèmes à base de connaissances sont développés pour exploiter une représentation de connaissances par le biais d'une logique de calcul implémentée sous la forme d'un algorithme. Les algorithmes manipulent les données afin de produire de nouvelles informations. Ils peuvent être appréhendés comme de réels processus de transformation et de réécriture du contenu des bases de connaissances basés sur les intentions de leur concepteurs et sur des savoirs et savoir-faire métier. L'algorithme et son fonctionnement représentent donc une information importante à véhiculer aux utilisateurs pour, par exemple, comprendre les résultats produits par ces systèmes à base de connaissances.

The image shows a software interface window titled "DOSSIER COMPOSANT". At the top, there are five tabs: "Informations", "Fonctionnement", "Programmes", "Compléments", and "Caractéristiques". The "Informations" tab is selected. The form contains several fields and buttons:

- Désignation:** A text input field containing "Dossier loi de Dirac".
- Description:** A large text area with a vertical scrollbar.
- Familles d'appartenances:** A text area with a vertical scrollbar and two buttons: "Ajouter" and "Retirer".
- Listes matérielles:** A text area with a vertical scrollbar.
- Listes fonctionnelles:** A text area with a vertical scrollbar.
- Matériel important pour la sécurité:** A dropdown menu currently showing "Non".
- Dossier de Conception:** A text area with a vertical scrollbar.
- Modes d'exploitation:** A text area with a vertical scrollbar and two buttons: "Ajouter" and "Retirer".

At the bottom right of the window, there are two buttons: "Valider" and "Annuler".

Figure 11 Un exemple de formulaire

Cependant, il nous faut constater que les systèmes informatiques se révèlent plus ou moins efficaces et adaptés si l'on en juge par les difficultés rencontrées lors de leur déploiement et de leur faible utilisation par les utilisateurs finaux. Ce constat ne semble pas lié à la pertinence ou l'intérêt de ces systèmes. Il semble plutôt que les moyens d'accès offerts, leurs interfaces restent souvent plus adaptées aux concepteurs, ou à des utilisateurs ayant une grande expertise

sur le système, au détriment des utilisateurs finaux. Il semble en effet que malgré le découpage en niveaux introduit par Newell (Newell 1981) les concepteurs se soient majoritairement concentrés sur les niveaux logiques et symboliques au détriment du Knowledge level (Rousseaux et Bouaziz 2006). Ce dernier renvoie pourtant à la problématique de la représentation des données à destination des utilisateurs humains et doit donc faire l'objet d'un intérêt particulier pour que les systèmes développés puissent être réellement valorisés.

### 2.3. Orientations de notre étude

Dans le cadre de cette étude, nous nous intéressons en particulier au problème de l'inscription d'une représentation des connaissances sur un support. Ce problème de l'inscription est un élément crucial dans le cadre de la capitalisation et de la restitution. L'orientation est donc de se focaliser sur trois aspects qui sont directement en relation avec les problématiques de capitalisation et de restitution :

- *Les supports,*
- *La définition des modalités d'inscription,*
- *Les outils pour lire et écrire.*

Le *support* joue un grand rôle à la fois dans la capitalisation et la restitution des connaissances. Il peut être papier ou numérique. On peut également selon un certain niveau d'abstraction considérer la dichotomie entre document numérique et système informatique comme correspondant à 2 types de support différents. Dans notre étude il s'agit donc d'étudier ce support qui va véhiculer l'inscription en dégageant en particulier les types de supports choisis et les raisons pour lesquels ils sont utilisés. Un deuxième élément est d'envisager la possibilité de produire une restitution multi support.

La *définition des modalités d'inscription* aborde tout aussi bien le problème de la forme de l'inscription que du fond, renvoyant à la sémantique qu'elle véhicule. Il s'agit en fait d'étudier en particulier la manière dont sont conçus les systèmes de visualisation d'informations. Un autre aspect correspond à l'analyse des méthodes de structuration lorsque l'on développe une ressource numérique. Plus généralement, il s'agit d'aborder le problème du développement de langage pouvant se matérialiser par une convention de mise en forme ou un langage de structuration d'informations standardisé.

Enfin, l'étude des *outils pour lire et écrire* est en relation avec les deux aspects précédents. L'inscription sur un support, qu'il soit numérique ou non, se fait via un artefact médiateur. Il s'agit donc d'identifier les outils qui sont choisis pour lire et écrire en dégagant les raisons de cette utilisation.

Notre étude selon ces 3 orientations doit nous amener à des propositions de supports, d'inscription et d'outils adaptés à la capitalisation et la restitution de connaissances. Avec cette orientation, notre approche théorique s'appuie sur trois axes en relation avec nos orientations :

- Ingénierie documentaire
- Ingénierie des interfaces homme machine
- Ingénierie de la forme

A travers l'ingénierie documentaire, nous abordons tout d'abord le document car il constitue l'objet de prédilection de la transmission du savoir. L'objet est de s'intéresser en particulier au concept, à l'évolution du document au vu de la numérisation et enfin à sa construction.

Nous poursuivons par une étude de l'interface homme machine en nous arrêtant en particulier sur la manière d'aborder leur conception. Il s'agit d'identifier quelques éléments de réponse permettant de faciliter leur appropriation.

Enfin, nous terminons par ce que nous nommons l'ingénierie de la forme regroupant à la fois le domaine de la visualisation d'information et de la sémiologie graphique. Il s'agit de nous intéresser à la manière de concevoir une forme visuelle pour des informations abstraites qui sont inscrites sur un support numérique.

## Partie II - Approche théorique

## Chapitre 3 - Ingénierie documentaire

### 1 - Du document au document numérique structuré

#### 1.1. Le document comme preuve

Le document est un objet, un concept, qui a traversé les siècles. Il constitue un « *objet unique, inscrit dans le temps et l'espace* » (Chabin 2004). Il renvoie classiquement à « *une trace volontaire résultant d'une activité humaine* » (Chabin 2004). Cette trace volontaire doit permettre de passer outre « le caractère éphémère de l'information » (Briet 1951) si elle n'est pas portée sur un support pérenne. Dans une vision classique, le document se matérialise souvent par une inscription sur un support : c'est le cas des documents papier. Il s'agit de la vision objet du document, évoquée par Paul OTLET (Buckland 1997), renvoyant « *aux manuscrit, imprimé, représentation graphique ou figurée, objet de collections etc...* » (El-Hachini 2005).

Cependant, la notion de document peut être étendue à d'autres formes notamment lorsqu'il est présenté comme « *une source d'information où l'on peut puiser des connaissances* » (Chabin 2004). Ainsi, Briet évoque « *l'antilope qui court dans les plaines d'Afrique ne peut être considérée comme un document... Mais si elle est capturée... et devient un objet d'études, on la considère alors comme un document. Elle devient une preuve physique* » (Briet 1951). Le document n'est alors plus limité à la forme traditionnelle de « *codex* » (le livre avec les pages reliées) et peut être associé à « *tout indice concret ou symbolique conservé ou enregistré aux fins de représenter, de reconstituer ou de prouver un phénomène ou physique ou intellectuel* » (Briet 1951).

Dans ce rapport au document, on souligne l'importance des attributs accompagnant le document : dans le cas de l'antilope, il s'agira d'une étiquette précisant qu'il s'agit d'une antilope. Ce sont ces attributs qui peuvent à la fois être source de connaissances pour le lecteur, mais également préciser sa légitimité dans sa valeur de preuve. La légitimité est à mettre en relation avec les notions d'autorité de l'information et de référence que véhicule la notion de document (Chabin 2004). L'autorité et la référence sont liés à l'auteur, son statut mais également sa localisation ainsi que l'institution qui le diffuse par exemple. Ceci souligne

l'importance du travail des sciences de l'information (Le\_Coadic 1994) et toute la tâche d'indexation du document qui va lui associer des attributs le caractérisant. Ainsi, la classification d'un document influencera nécessairement ce document.

Dans notre étude, nous désignons le document comme « *production sémiotique transcrite ou enregistrée sur un support pérenne qui est équipé d'attributs spécifiques visant à faciliter les pratiques liées à son exploitation ultérieure dans le cadre de la préservation de transactions communicationnelles distribuées. Ces attributs doivent permettre au document de circuler à travers l'espace, le temps, les communautés* » (Zacklad 2004).

## 1.2. Le document numérique

La numérisation a eu un impact important sur la notion de document (Pédauque 2006). Au travers de cette numérisation du support et de l'inscription, le document se dématérialise. Il devient un enregistrement numérique qui peut être aisément modifié et réutilisé. Le document n'est plus caractérisé par sa finitude mais par une ouverture, une facilité de recombinaison de la forme et du fond. Ainsi, le document numérique peut être « *défini comme une représentation numérique d'une preuve, qui doit pouvoir être réutilisée (reused) dans un autre processus de traitement* » (Chabin 2004). Les frontières du document sont alors largement déplacées du fait de la facilité de réutilisation du contenu par un homme ou un ordinateur.

Le document numérique est largement influencé par l'évolution technologique et la création de formats numériques. Bien que le numérique implique une dématérialisation, ces formats se sont inspirés de la réalité et en particulier du support final visé. Il semble toutefois que l'évolution des formats tend vers « *une abstraction croissante vis-à-vis du médium physique* » (Lux-Pogodalla et Vion-Dury 2004). Ainsi, cette abstraction peut être illustrée par les « *formes documentaires* » suivantes d'abstractions croissantes (Lux-Pogodalla et Vion-Dury 2004) :

- Images et sons digitaux,
- Descriptions de page,
- Contenus orientés présentation,
- Documents structurés.

Les *images et sons digitaux* (TIFF, GIF, BMP...) se caractérisent par une représentation numérique des pixels ou échantillons sonores. La manipulation de ce type de document pourra se faire ou en s'appuyant sur des techniques issues du traitement du signal ou par une pratique manuelle via un éditeur spécialisé.

Les *descriptions de page* (PS, PDF) se basent sur une métaphore du document papier. La représentation numérique se fonde sur une description de la manière dont il serait construit sur un support papier. Le format permet ainsi de préciser l'ensemble des attributs inhérents à une page (format, marge,...), le positionnement d'éléments dans cette page (blocs de texte, images...) ou encore la mise en forme matérielle du texte (gras, italique,...). La modification du document devient alors une opération complexe du fait de la forte imbrication du contenu, de la mise en forme, mais également de la définition du medium dans le même fichier.

Les *contenus orientés présentation* sont en relation avec de nombreuses solutions caractérisées par des formats propriétaires. L'exemple caractéristique est le traitement de texte Microsoft Word. La représentation numérique mélange à la fois le contenu, les informations de style, les graphiques ou encore les informations de structure logique. La manipulation des fichiers est possible via une opération d'édition à condition d'être en possession d'un logiciel auteur permettant de lire le format propriétaire : il s'agit souvent du logiciel ayant permis de le créer.

Les *documents structurés* (SGML, XML) représentent une forme d'abstraction très importante puisqu'ils se caractérisent par une dissociation très franche entre le contenu et la mise en forme. Ils bénéficient d'une indépendance par rapport aux logiciels auteur ayant permis de les créer ce qui leur confère à la fois une facilité de traitement et une certaine pérennité. Ils constituent une forme abstraite très aboutie du document.

### 1.3. Le document numérique structuré

#### 1.3.a Définition

Le document numérique structuré ne peut être réduit à la simple évocation de l'utilisation des technologies XML (eXtended Markup Language). Pour le définir, nous prenons les éléments proposés par (Lux-Pogodalla et Vion-Dury 2004) : « *les entités élémentaires sont des éléments structuraux (ou logiques) dont la signification est paramétrable. Les règles d'organisation logique sont explicites et modifiables séparément des instances documentaires*

(schémas documentaires). *La validité structurelle d'un document par rapport à un schéma peut être vérifiée via des algorithmes génériques. ».*

Le document structuré est alors une structure dont les éléments (les unités élémentaires) vont stocker des contenus textuels. Cette structure sera classiquement hiérarchique. Ce mode d'organisation de l'information définit alors une réelle syntaxe dans la construction du document. Ainsi, les différents éléments structuraux qui délimitent les entités élémentaires d'un document associés à leur règle d'organisation constituent un réel langage : un contrat d'écriture et de lecture dans le document. En fonction de la technologie utilisée, les mots du langage de structuration sont plus ou moins contraints. Ainsi, le langage HTML (Hypertext Markup Language) définit une structuration de l'information dont les balises sont prédéterminées et répondent à une mise en forme standard : un exemple est présenté sur la Figure 12. Il en est de même lors de la création d'un document en Tex (cf. Figure 14 et Figure 15), standard développé par Donald Knuth (Knuth 1984).

```
<TR>
  <TD>Nom</TD>
  <TD>
    <INPUT type=text name="nom" >
  </TD>
</TR>
<TR>
  <TD>Prénom</TD>
  <TD>
    <INPUT type=text name="prenom" >
  </TD>
</TR>
<TR>
  <TD>Sexe</TD>
  <TD>
    Homme : <INPUT type=radio name="sexe" value="M">
    <br>Femme : <INPUT type=radio name="sexe" value="F">
  </TD>
</TR>
```

Figure 12 Exemple de code html et de sa visualisation

Le standard XML (eXtended Markup Language) pour sa part ne constitue pas en soi un langage « *mais plus exactement un ensemble de règles permettant la création de langage de balisage* » (Ray 2001). Ainsi, en s'appuyant sur XML, les éléments de structuration de l'information ne sont pas prédéfinis (cf. Figure 13). La seule contrainte est de produire des documents bien formés, c'est à dire en se conformant aux règles lexicales et syntaxiques du standard XML (fermeture des balises ouvertes, pas de chevauchement de balises, etc...).

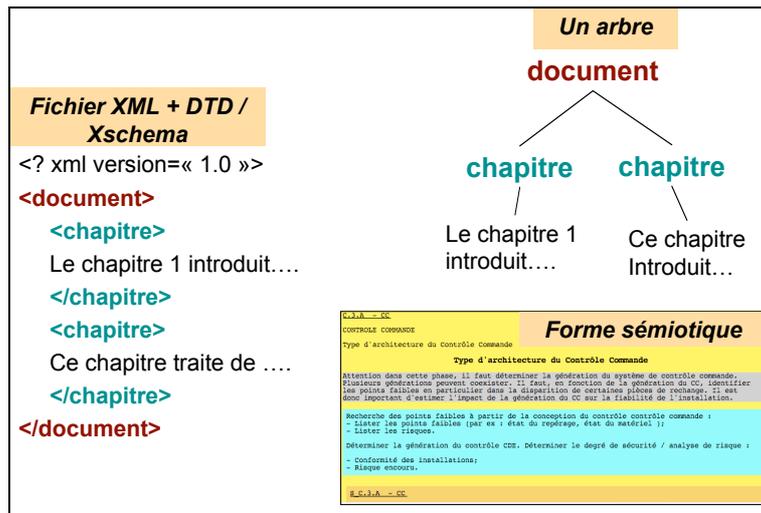


Figure 13 – Document structuré en XML sous différentes formes

### 1.3.b Notion de langage de structuration

Ainsi, le langage de structuration (concrétisé par le balisage) peut être construit selon 2 points de vue (Marcoux 2006) :

- Le balisage logique,
- Le balisage généraliste

Le *balisage logique* est défini pour traduire l'enchaînement logique du document, à l'image de la structuration TEX (Knuth 1968) (cf. Figure 14), mais en conservant une indépendance au regard de la mise en forme matérielle que prendra le contenu. Il s'agira par exemple d'insérer au sein du contenu des éléments structurant délimitant les titres, les parties, les sections, etc... La mise en forme du document pourra ensuite être faite au travers de la mise en relation des éléments structurants avec une feuilles de style. La feuille de style constitue une réelle légende qui spécifie la mise en forme associée à chaque brique élémentaire du document.

```

\documentclass{article}           % Type de document
\title{Document Latex simple}    % Titre du document
\author{Patrice Guay}           % Auteur
\date{15/08/2004}               % Date de création
\begin{document}                 % Début du document
\maketitle{}                     % Génération du titre
\section{Introduction}           % Section #1
Ce document est un exemple de fichier
source Latex.

```

```

\section{Developpement}                % Section #2
Il s'agit d'un document de type "article".
Son titre est "Document Latex simple" et
il contient trois sections : "Introduction",
"Developpement" et "Conclusion".
\section{Conclusion}                  % Section #3
Ceci termine notre premier exemple de
fichier source Latex.
\end{document}                        % Fin du document

```

Figure 14 Exemple de document structuré en Tex. (source :

[http://www.nanotechnologies.qc.ca/propos/linux/documentation/Redaction\\_documents\\_Latex/ch-structure\\_latex.html](http://www.nanotechnologies.qc.ca/propos/linux/documentation/Redaction_documents_Latex/ch-structure_latex.html))

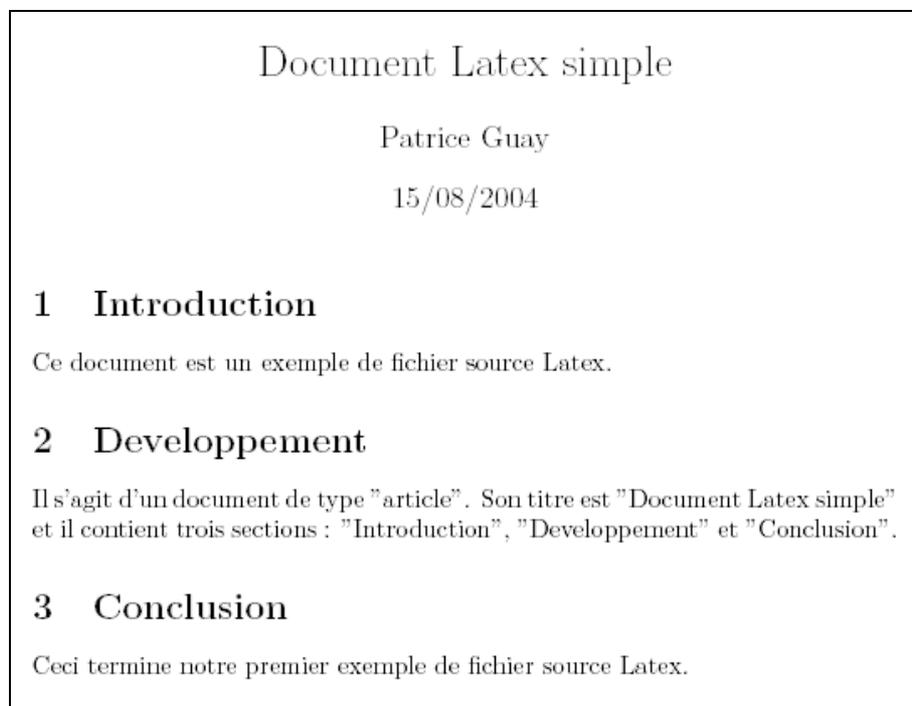


Figure 15 Document Tex mis en forme

Le *balisage généraliste* se positionne autrement puisque les éléments structurants insérés dans le document vont constituer un métalangage. Ils pourront alors caractériser plus sémantiquement le contenu sans être trop influencés par la métaphore du document papier et par exemple l'enchaînement linéaire du contenu imposé par la succession des pages. Le balisage généraliste tendra alors à définir des langages de structuration propre à un domaine ou une classe de document. Pour un courrier, on identifiera l'expéditeur, le destinataire, la date, le sujet ou le corps (cf. Figure 16). Pour un formulaire de commande, les éléments structurants correspondront au nom, prénom, date, adresse de livraison, articles commandés et

moyen de paiement. Le balisage généraliste traduit une forme d'indexation très fine du contenu.

```
<?xml version="1.0" encoding="UTF-8"?>
<courrier>
  <expediteur>Wallace</expediteur> <destinataire>Gromit</destinataire>
  <date>
    <jour>25</jour> <mois>octobre</mois> <année>2007</année>
  </date>
  < sujet> Manque de Cheddar </sujet>
  <corps>
Gromit nous sommes à cours de Cheddar. Or, la vie est impossible sans
Cheddar.
A bien y regarder, la lune a de jolis trous, comme le délicieux fromage.
Nous pourrions partir en vacances sur la Lune.
  </corps>
</courrier>
```

Figure 16 Exemples de balisages généralistes

### 1.3.c Modèle de document, document bien formé et document valide

Que le balisage soit logique ou généraliste, les documents structurés sont fondés sur la définition d'un ensemble d'éléments structuraux qui vont constituer les mots d'un langage dédié à la structuration du contenu. Ce langage peut être ouvert ou fermé. Lorsque le langage est ouvert, il est possible de définir de nouveaux mots – c'est-à-dire de nouveaux éléments structuraux - qui vont enrichir le langage et permettre de délimiter de nouveaux contenus. Bien que le langage soit ouvert, la construction du document doit en général répondre à certain nombre de règles d'écriture qui vont par exemple permettre d'identifier les éléments structuraux par rapport au contenu.

Ce point permet d'aborder la notion de *document bien formé*. Un document bien formé est un document structuré qui va se conformer à des règles élémentaires de définition des balises et de structuration. Dans le langage XML, il s'agira (cf. Figure 17) :

- de définir les balises sous cette forme « <nom\_balise> »

- qu'une balise ouverte soit fermée par une balise « `</nom_balise>` »
- qu'il n'y ait pas de chevauchement entre des balises différentes

Dans le cas d'un langage fermé, les éléments structuraux sont définis et ne supportent pas l'ajout de nouveaux mots. Un exemple est le langage HTML. Les langages fermés permettent d'introduire la notion de document valide. Le caractère fermé permet de définir de manière précise les règles lexicales et syntaxique d'un langage de structuration. Elles spécifient les éléments structurants disponibles et utilisables ainsi que leur organisation les uns par rapport aux autres. Ces règles peuvent être traduites sous la forme d'un *modèle de document* (cf. Figure 18) qui permettra alors de s'assurer qu'un document structuré est en accord ou non avec le langage défini. Il s'agit là de la validation introduisant la notion de *document valide* (cf. Figure 18). Notons au passage qu'un document valide est nécessairement bien formé.

Le document doit avoir exactement un élément de niveau le plus élevé appelé : élément document ou élément racine. Tous les autres éléments doivent être logés entre ces marqueurs.

```
1 <racine>
2   <autres_balises>...</autres_balises>
3   ...
4 </racine>
```

*Les éléments doivent être correctement emboîtés. Les éléments doivent être imbriqués, mais ne doivent pas se recouvrir.*

```
1 <livre><auteur>...</auteur></livre> // correctement emboîté
2 <livre><auteur>...</livre></auteur> // attention, écriture incohérente
```

*Chaque élément doit avoir à la fois un marqueur d'ouverture et un marqueur de fermeture. Si nous avons besoin d'un élément vide et que nous désirons traiter directement avec les attributs, nous pouvons utiliser un raccourci.*

```
1 <balise> // marqueur d'ouverture .....</balise> // marqueur de fermeture
2 <vide></vide> ----> <vide />
```

*Les noms d'éléments sont sensibles à la casse.*

```
1 <TITRE> .....</titre> // attention, ces balises non rien à voir entre elles
```

Figure 17 Quelques exemples de règles devant être respectées pour qu'un document XML soit bien formé

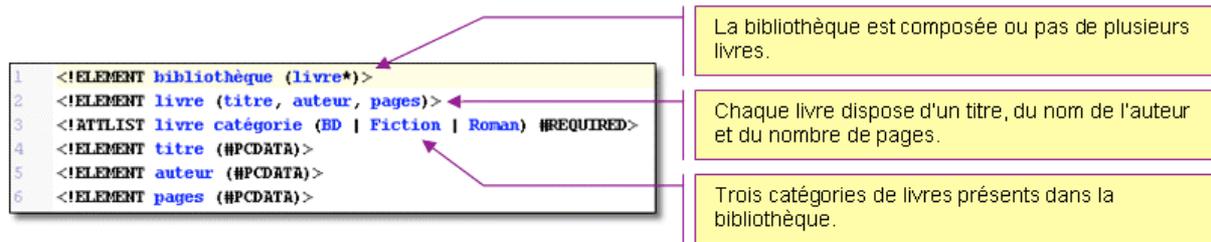


Figure 18 Illustration d'un modèle de document sous la forme d'une DTD (source : <http://pagesperso-orange.fr/emmanuel.remy/XML/DTD/DTD.htm> )

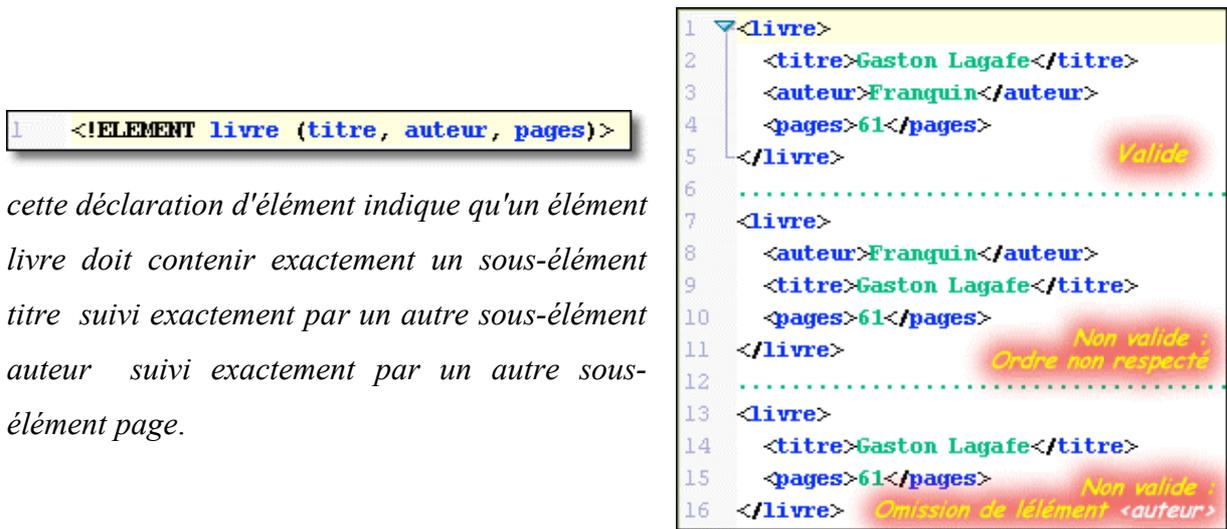


Figure 19 Illustration de la validation d'un document XML (source : <http://pagesperso-orange.fr/emmanuel.remy/XML/DTD/DTD.htm> )

La définition de (Lux-Pogodalla et Vion-Dury 2004) évoquée précédemment met l'accent sur l'importance d'explicitier les règles d'organisation logiques des entités élémentaires et le fait qu'elles doivent pouvoir être validées de manière algorithmique dans le cadre de documents structurés numériques. Ainsi, les modèles de document qui sont typiquement implémentés sous la forme d'une DTD (Document Type Definition) ou d'un Xschema (XML schema) lorsqu'on exploite les technologies XML, représentent une source d'informations essentielle. En effet, ils régissent le contrat de lecture, d'écriture et de validation des documents. Ils permettent de valider le document afin par exemple de vérifier son origine et son autorité.

## 2 - La réutilisation face au concept de document

L'analyse interdisciplinaire synthétisée dans (Pédaque 2006) identifie trois dimensions d'analyse pour le document que sont la *forme*, le *signe* et le *medium*. La *forme* appréhende le

document comme un *objet matériel ou immatériel* qui va présenter une visualisation de l'information. Le *signe* renvoie plutôt au document comme *objet signifiant*. La dimension *medium* appréhende le document comme un « *vecteur de communication* » et fait référence au *processus social de mise en document*. Ces dimensions ne peuvent être analysées de manière indépendante car elles s'influencent et se recoupent. Le document constitue en effet un « tout » (Rastier 1987) qui représente plus que la somme de ces parties.

## 2.1. Les reformulations du numérique

Nous proposons ici une vision synthétique de l'analyse interdisciplinaire du concept de document lors de son passage au numérique, proposée dans (Pédauque 2006), sous les trois dimensions que sont la forme, le signe et le medium.

### 2.1.a Le document comme forme

Le document traditionnel comme forme est appréhendé comme un objet de communication et de diffusion de l'information. La forme se traduit en particulier par la mise en forme matérielle du texte ou des illustrations contenus dans le document. Ces règles de mise en forme, qui révèlent le style visuel du document, fixent un contrat de lecture. Le contrat de lecture renvoie aux modalités de lecture d'une inscription sur un support. Il est construit par l'auteur qui doit définir la relation entre un support et le lectorat ciblé. Le document traditionnel peut alors être défini par l'équation :

$$\textit{Document traditionnel} = \textit{support} + \textit{inscription}$$

Le document est ici caractérisé par un support tangible (le papier par exemple) et une forme d'inscription. Cette inscription, qui traduit à la fois le contenu et la visualisation sous laquelle on la perçoit, répond à un contrat de lecture qui pour le document traditionnel est identique au contrat d'écriture. Ce contrat de lecture sera plus ou moins explicité, ce qui pourra le rendre plus ou moins difficile à appréhender par les lecteurs. Le numérique introduit une certaine dématérialisation de l'inscription ce qui complique le rapport au support. Il devient difficile à identifier car ce support correspond « *au fichier, à l'outil matériel qui l'héberge ou encore à la surface de l'écran* » (Pédauque 2006). L'inscription est également impactée par la numérisation. Elle est en effet très cadrée et « *relève du codage, une opération familière des informaticiens* » (Pédauque 2006). Elle suit alors des règles formelles rigides, correspondant alors à des données enregistrées en mémoire sous une forme permettant d'y accéder. Cette forme doit alors répondre à une structure pouvant être modélisée et manipulée par la machine

indépendamment du support. Ainsi, l'équation précédente se transforme dans le monde numérique en :

$$\textit{Document numérique} = \textit{structure} + \textit{données}.$$

Le fait de se dégager de la notion d'inscription écarte de l'équation la mise en forme. Ce n'est plus tant la mise en forme qui définit le contrat de lecture, mais la structure qui offre une lisibilité partagée et prédéterminée entre le producteur (ex : un informaticien) du contenu et le lecteur qui pourra se révéler être numérique (un programme par exemple). Les fichiers CSV, présentés dans le Chapitre 1 -2.2.b , illustrent cette situation. En effet, la mise en forme va traduire de manière formelle et déterministe une structure identifiable par l'humain ou la machine : le contrat de lecture est basé sur la définition d'une codification précise, c'est-à-dire un langage monosémique. De même, les bases de données sont caractérisées par leur modèle relationnel. Ce modèle définit la structure et le contrat de lecture et d'écriture dans la base.

Les évolutions dans le domaine du document numérique, et notamment toute la philosophie accompagnant les technologies de structuration de contenu, en particulier XML, redéfinissent cette dernière équation pour aboutir à :

$$\textit{Document XML} = \textit{données structurées} + \textit{mise en forme}.$$

Le document XML doit être lu comme document structuré. Il se caractérise par une structure logique permettant d'identifier les entités élémentaires du document. Ces entités élémentaires vont définir un découpage et une organisation de l'information analysable par l'ordinateur. La mise en forme (le « style ») est alors définie indépendamment des données structurées, elle devient modulable et prend donc une place importante dans le sens du document. Il est en effet possible de décliner la visualisation du document sous de nombreuses formes. Le « style » traduit alors un mode de représentation, une objectivation (Bachimont 2004) des données structurées, ce qui « *délaisse la forme comme dimension primordiale du document* » (Pédauque 2006) du fait de la facilité de recomposition.

### *2.1.b Le document comme signe*

Le document traditionnel comme signe considère le document comme un objet signifiant. Cet objet signifiant s'appuie majoritairement sur le sens véhiculé par le contenu et donc l'inscription. Une inscription est ainsi porteuse de sens, le lecteur pouvant accéder à ce sens

s'il est en mesure de « déchiffrer » l'inscription selon le contrat de lecture propre au document. Le document traditionnel peut alors être défini comme :

$$\textit{Document} = \textit{inscription} + \textit{sens}$$

Dans le monde numérique, on peut modifier l'équation et aboutir à :

$$\textit{Document numérique} = \textit{texte informé} + \textit{connaissances}$$

La substitution « d'inscription » par « texte informé » tend à signifier que le texte a été soumis ou pourrait être soumis à un traitement permettant d'en repérer les unités d'information. C'est en particulier le cas des documents structurés qui intègrent des éléments structuraux, un balisage, permettant de préciser le découpage du document, en délimitant des fragments, selon un point de vue logique ou généraliste. Plus généralement, il s'agit d'introduire dans le contenu ou d'associer au texte des métadonnées qui vont faciliter le traitement, le repérage et la localisation des documents par exemple.

Le remplacement de « sens » par « connaissances » voudrait introduire la notion de personnalisation pour un lecteur ou un usager donné. Il ne s'agit plus uniquement de mettre à disposition les documents, mais de fournir une réponse adaptée aux demandes, aux questions de nos usagers. Ceci débute en premier lieu par le fait de fournir le bon document à un lecteur et, pourquoi pas, d'en adapter l'accès ou la visualisation pour une activité précise.

Dans la perspective de construire des documents structurés, il s'agit de valoriser au mieux cette ressource manipulable par l'ordinateur. Les unités d'informations étant repérées, il est possible de caractériser avec précision les informations repérées. Cela peut rendre plus pertinent les traitements opérés sur les documents et leur contenu. Ainsi, l'important travail sur le Web Sémantique tendrait à aboutir à l'équation :

$$\textit{Document Web Sémantique} = \textit{texte informé} + \textit{ontologies}$$

Dans le document Web Sémantique, l'idée est de compléter le texte informé de métadonnées permettant de préciser l'espace sémantique et le contexte dans lequel des unités d'informations délimitées doivent être interprétées. Ces métadonnées peuvent être modélisées sous la forme d'ontologies. Elles auront l'ambition, par exemple, d'assister un moteur de recherche pour en améliorer les résultats. Les ontologies sémiotiques assisteront un lecteur humain dans son accès à l'information et son interprétation du contenu. Un autre dispositif,

pour assister un utilisateur humain, pourrait être les annotations (Lortal et al. 2007; Zacklad 2007) qui vont permettre d'enrichir le contenu d'éléments connexes.

### *2.1.c Le document comme medium*

En analysant le document traditionnel dans sa dimension medium, on s'intéresse au processus social de mise en document. Le statut de document s'acquerrait alors sous deux conditions :

- Dépasser la communication intime (entre quelques personnes privées)
- Dépasser l'éphémère

L'inscription doit donc dépasser la communication intime, la diffusion ne devant pas être limitée « entre quelques personnes privées ». Briet (Briet 1951) met l'accent sur le fait que « *la documentation secrète est une injure à la documentation* ». Ainsi, le document doit être rendu disponible et faire l'objet d'un repérage pour acquérir une légitimité. La légitimité dans le document traditionnel est également liée au temps. L'inscription ne doit pas être éphémère et donc dépasser « le moment de son énonciation ». Le document représente une inscription sur un support pérenne. Ceci lui attribue une certaine légitimité. En effet, il pourra alors être transmis et traverser le temps au prix d'un effort sur ses modalités de conservation. Le document peut alors être défini par l'équation :

$$\textit{Document} = \textit{inscription} + \textit{légitimité}$$

La légitimité associée à l'inscription explique bien l'utilisation du document comme objet de transmission du savoir – les racines latines du mot « document » renvoient à l'activité d'enseignement - mais également de « *régulation des sociétés humaines* ». Avec la numérisation du support, les documents sont de plus en plus faciles à éditer, reprendre et transmettre. Ainsi, on observe une croissance importante de la mise en document des informations liées à cette facilité de création. Ainsi, dans les entreprises et les organisations, le document constitue un moyen à la fois de coordination et de supervision des activités. La coordination est en relation avec la dimension coopérative du travail qui nécessite en particulier de médiatiser une représentation partagée. La supervision des activités est à destination des managers, par exemple, qui, en observant la publication des documents, peuvent obtenir une image de l'avancement des activités. Cet avancement devient observable par l'intégration de démarches qualité qui définissent, sous la forme de procédures, une succession de documents types à produire. Ces procédures, intégrées à des systèmes de

workflow, « *amplifient la production de documents* ». Dans son passage au numérique, l'équation précédente devient alors :

$$\textit{Document numérique} = \textit{texte} + \textit{procédure}$$

Le document numérique correspond alors à une trace d'une relation sociale ou d'une pratique pouvant être calculée au travers d'un processus informatique notamment lorsque les documents deviennent très structurés. La légitimité dans cette équation est toujours présente, mais se traduit par la procédure de création ou encore de diffusion du document dans l'organisation. Ainsi, les statuts courants associés au document tel que « en cours de rédaction », « pré diffusion » ou « diffusion » auront un impact sur la légitimité des documents lorsque les lecteurs y accéderont. Les modalités d'accès prennent également une grande importance lors du passage au numérique, le document ne pouvant être vu comme un élément isolé. Son positionnement dans un plan de classification ou l'emplacement sur lequel il est sauvegardé, influence, le lecteur. Ainsi, pour les documents Web par exemple, le positionnement sur la toile, traduit par l'URL (Uniform Resource Locator) peut traduire une certaine légitimité au regard de l'information fournie. Le document WEB se traduit sous l'équation :

$$\textit{Document Web} = \textit{publication} + \textit{accès repéré}$$

La légitimité s'est alors transformée par un accès repéré. Le medium ne se limite plus au support sur lequel est porté l'inscription tel que le papier ou un support numérique. Le chemin d'accès au document est associé au « tout » qu'il constitue et lui confère une légitimité.

## 2.2. La réutilisation face au concept de document

Du document traditionnel au document numérique, la section précédente illustre une évolution au regard des dimensions forme, signe et medium. Le document qui constituait un objet tangible devient manipulable de manière aisée sur l'ensemble des dimensions. Tout ceci s'accompagne d'une facilité d'édition de plus en plus déconcertante, rendant le document de plus en plus réutilisable.

Au travers de la réutilisation, on s'écarte sensiblement de la notion de « tout » que constitue le document. Ainsi, « *le numérique déconstruit un document en : une ressource enregistrée + une reconstruction calculée dynamiquement* » (Bachimont et Crozat 2004). « *La déconstruction conduit à une perte d'objectivité et à une dilution du document* » (Bachimont

et Crozat 2004). Cette dilution est liée au calcul dynamique et donc à la définition des algorithmes de transformation des documents. Ces algorithmes constituent de réels processus de réécriture, pouvant impacter tout aussi bien sur la forme que sur le fond. Ils prennent alors une importance et une responsabilité énorme dans la production et la diffusion de l'information. En effet, la notion de ressource enregistrée et la facilité de recomposition des contenus, que cela soit sur la dimension signe ou forme, pose le problème de voir le document comme un « *lego* » (Cotte et Desprès-Lonnet 2004). Se pose alors la question de « *l'autorialité* » et du travail éditorial, un exemple caractéristique étant la pertinence de construction, correspondant plutôt à un calcul, de revue de presse automatique (Cotte et Desprès-Lonnet 2004) telle que Google News.

En effet, la réalisation d'un document correspond le plus souvent à une activité collective de rédaction mettant en jeu un réseau d'acteurs qui vont plus ou moins contribuer à sa définition que cela soit en terme de *forme* (travail éditorial) ou en terme de *signe* (travail autorial). Ainsi, un document, avant de constituer un « tout », va passer par tout un ensemble de versions et constituer un objet relativement instable, mais qui prend déjà un rôle important dans les activités et les organisations qu'il instrumente. A ce stade, nous définissons ces documents, comme des *documents pour l'action*.

### 3 - Le document pour l'action

#### 3.1. Définition

Le document pour l'action (DopA) (Zacklad 2004) permet de redéfinir le concept de document et d'appréhender autrement leurs contenus qui relèvent de moins en moins de la catégorie du texte classique. En insistant sur la dimension collective de l'activité rédactionnelle, le DopA permet d'analyser les documents comme relevant de processus de communication pour partie différés, au sens des processus asynchrones décrits dans le champ du CSCW, entre des producteurs et des récepteurs liés par des intérêts communs. Alors que la conceptualisation en terme d'hypermédia visait surtout à rendre compte des nouvelles pratiques de lecture associées aux hypertextes, le DopA vise à rendre compte des processus « d'hyper-rédaction » associés aux documents numérisés. Le DopA se définit alors comme « *un ensemble de fragments portés par des auteurs divers dont le contenu final reste largement indéterminé alors même que sa circulation rapide lui fait déjà jouer un rôle majeur d'information, d'aide à la décision et de preuve* » (Zacklad 2004).

Notons bien que cette définition vient pour partie en contradiction avec le document vu comme medium évoqué précédemment. En effet, le DopA peut d'une part renvoyer à une communication « intime » : dans une situation de travail, la feuille de brouillon exploitée par une personne dans une situation de résolution de problème peut prendre le statut de DopA. D'autre part, il n'exclut pas le caractère « éphémère », ou tout au moins temporaire, ce qui vient en contradiction avec les attributs de légitimité. Ainsi, à l'image de ce qu'évoque Morand (Morand 2004) concernant l'utilisation des diagrammes en conception logicielle, le DopA peut être vu comme un support d'espace de résolution de problème : les instances stables ou même finales peuvent être d'une importance secondaire par rapport aux évolutions entre chacune d'elles qui reflètent d'avantage l'activité associée à la construction du signifiant. Ainsi, le DopA ne vient pas en contradiction de la notion classique de Document mais est plutôt complémentaire. Le DopA offre un cadre d'analyse couvrant la construction du document.

### 3.2. Construction des documents

La numérisation du document souligne l'importance du concept de DopA. En effet, au vu de l'accélération de la production documentaire, lié à la facilité de réutilisation de la forme et du contenu, le document dans sa construction rend compte d'une grande instabilité. Le document passe d'une version à une autre avec une simplicité déconcertante, pouvant conserver un historique des modifications et être enrichi d'annotations via les mécanismes de révision des traitements de texte par exemple. Nous proposons d'aborder ici la problématique de construction des documents en commençant par introduire la notion d'architexte. Nous évoquerons ensuite la problématique de construction de deux types de documents numériques :

- Les hypertextes
- Les documents structurés

#### 3.2.a *Les architextes*

La réutilisation des documents et de leur contenu, pour partie, ne peut être dissociée des logiciels auteur qui permettent de produire ces documents. Ainsi, dans (Jeanneret et Tardy 2007) la notion d'*architexte* est introduite pour y faire référence. Un architexte est défini comme un « *logiciel considéré dans sa capacité à mettre en forme et conditionner l'écriture. L'architexte (traitement de texte, navigateur, logiciel de présentation, etc.) procède à une*

---

*écriture de l'écriture : il relève de la production de formes qui se situent en amont de l'acte d'écriture-lecture et en déterminent les conditions de possibilité (du grec, origine et pouvoir) » (Jeanneret et Tardy 2007).* Ces architextes constituent des interfaces de création et de diffusion des documents qui vont à la fois contraindre l'écriture par leur possibilités, nécessairement limitées, mais également faire l'objet d'utilisations inédites par les utilisateurs qui les adaptent à la tâche, les transformant ainsi en instrument (Béguin et Rabardel 2000).

Les contraintes d'écriture, imposées par les architextes, sont liées aux moyens graphiques et interactifs mis à disposition dans l'activité d'écriture et d'édition. Les moyens graphiques renvoient aux possibilités de gestion d'espace dans l'écran ou la page, à l'intégration d'icônes ou d'images ou encore aux capacités typographiques permettant de mettre en forme le texte. Les moyens interactifs sont par exemple la possibilité d'intégrer des liens hypertextes. Il est ainsi possible de dégager des pratiques récurrentes et spécifiques d'écriture, liées à chaque architexte (Jeanneret et Tardy 2007). Si l'on prend l'exemple de Powerpoint, il s'agit d'un mode d'écriture séquentialisé dans lequel on scénarise la diffusion et la présentation du contenu au lecteur : on évoque « *une représentation de l'oral dans l'écrit* » (Jeanneret et Tardy 2007). Dans (Gaglio et al. 2006), on présente ce mode de diffusion de l'information comme un « écrit théâtral ». L'architexte influence alors le genre d'écriture en particulier sous l'aspect de la forme.

### 3.2.b L'hypertexte

Le support numérique permet d'intégrer au document des possibilités peu envisageables par le document traditionnel et en particulier l'interactivité. Cette interactivité passe en particulier par l'intégration de liens hypertexte dans les documents qui peuvent être intra documentaire ou inter-documentaires. Les liens intra-documentaire permettent d'offrir des parcours de lecture inédits et passer outre la lecture linéaire classique. Les liens inter-documentaires permettent à la fois de positionner un document par rapport aux autres et de l'enrichir par la même occasion. La définition de liens permet d'aboutir à un nouveau type de document : les documents hypertextes.

L'hypertexte ainsi que l'hypermedia avec l'introduction d'images, de sons et de vidéo, sont des termes inventés par Ted Nelson en 1968. Le concept d'hypertexte précède le document numérique puisque sa paternité est attribué à V. Bush, T. Nelson reprenant et étendant ses approches, qui en 1945 introduit le Memex comme une mémoire externe basée sur la mise en œuvre de mots clés, de références, d'indexation, de liens et d'annotation. Il s'agit alors du

premier hypertexte basé sur des microfilms. V. Bush veut ainsi au travers de ce mode de représentation s'inspirer de la pensée humaine, sans volonté de vouloir l'égaliser, celle-ci opérant par association.

L'hypertexte se définit alors comme un ensemble de documents liés entre eux par des hyperliens. Il présuppose un nouveau genre d'écriture qu'il faut manier avec précaution. L'intégration de liens suppose alors pour partie une lecture navigationnelle pouvant aboutir à des problèmes de désorientation et de surcharge cognitive (Ertzscheid 2002), que cela soit dans l'écriture ou la lecture du document. Il s'agit donc d'instrumenter au mieux la construction de ces documents via des architextes adaptés. Laissant de côté les grands acteurs commerciaux, de nombreuses études tendent à produire des solutions d'édition hypertexte.

HyWebMap représente un exemple très aboutit qui, parmi d'autres fonctionnalités, assiste l'édition de ce qui est dénommé *réseau virtuel* ou *espace de connaissances personnel* (Saleh 2005). Un réseau virtuel correspond alors à un hypertexte ou un hypermedia pouvant à la fois lier « *des ressources documentaires exogènes puisées sur Internet et/ou au sein d'un Intranet [...] et également à partir de ressources personnelles* » (Saleh 2005). Ainsi, HyWebMap se positionne tout à fait comme un outil de réutilisation de contenus, puisque la définition de réseaux virtuels revient à réorganiser l'espace d'informations pour aboutir à la création d'un nouvel hypertexte enrichi, par exemple, d'annotations.

### 3.2.c Les documents structurés

L'hypertexte s'est largement développé en s'appuyant sur la technologie HTML produisant ainsi des documents structurés. Le caractère ouvert et standardisé du format a nécessairement contribué au développement du Web. La structuration des documents permet ainsi de cibler avec précision au sein d'un document le fragment pointé par un lien. L'édition de ce type de document reste encore une problématique importante. En effet, il est aujourd'hui aisé d'obtenir un éditeur permettant de produire des documents structurés en HTML. Ces éditeurs proposent alors une solution adaptée à une structuration logique de l'information ou l'ensemble du langage est connu et prédéterminé. Les éléments de langage répondent à une sémantique et il est alors possible d'y associer une mise en forme standard. Les solutions commerciales sont nombreuses. Nous nous limiterons ici à évoquer la solution libre Amaya (Quint et Vatton 2004). Cet éditeur se focalise sur l'écriture de document en XHTML et offre une solution permettant de construire des documents mixant différents langages de structuration sans exiger une bonne connaissance de ces derniers.

L'édition de documents structurés basée sur des langages inédits pose le problème à la fois de la sémantique des éléments structurants et de leur représentation. Le document structuré peut ainsi tout à la fois mixer un balisage logique et un balisage généraliste. Intuitivement, le balisage logique qui définit un découpage du document pourra être traduit par une hiérarchie de titres, alors que le contenu des balises généralistes pourra faire l'objet d'une mise en valeur. Il n'en reste pas moins que ce processus de traduction doit faire l'objet d'une définition qui précisera la mise en forme sous la forme d'une feuille de style par exemple. L'édition de documents structurés, en particulier en XML, reste encore peu instrumentée. Bien entendu, il existe de nombreux éditeurs en code source, mais qui ne peuvent que difficilement être mis entre les mains d'utilisateurs non informaticiens. Les grands acteurs du marché, comme la société Altova, qui propose XML Spy, offre une solution tout en un qui se révèle coûteuse en ressource et complexe, ce qui en réserve l'utilisation à des spécialistes. Amaya offre un début de solution puisqu'il se focalise sur l'édition de document XHTML, qui, s'il ne propose pas la flexibilité d'un métalangage comme XML, permet tout de même de définir balisage logique et généraliste.

L'édition des documents structurés ne se limite pas à l'intégration d'éléments structurants dans un contenu. La définition de la sémantique inhérente aux balises est essentielle, l'idée étant par exemple de pouvoir dissocier les balises logiques, des balisages généralistes et pour ce dernier type qualifier avec précision le contenu pour par exemple en adapter la représentation. La documentation des éléments structurants constitue donc un enjeu majeur permettant de valoriser au mieux cette forme documentaire. Dans cette direction, Marcoux propose une approche basée sur le langage naturel pour la construction de documents structurés en XML (Marcoux et Rémillard 2001; Marcoux 2006). Dans (Nanard et Nanard 2005), on propose de caractériser l'intention liée à la définition d'une balise dans un contenu ce qui peut alors permettre de définir des documents hypermedias adaptatifs.

#### 4 - Du document à l'interface homme machine

Comme nous avons pu le voir le document a subi une importante mutation dans son passage au numérique. Dans ces versions les plus évoluées, le document numérique, en particulier lorsqu'il est très structuré, offre un éventail de possibilités le rendant très malléable sur le plan de la forme et du fond. Le document numérique structuré permet ainsi d'abstraire les structures (Bachimont 2004) :

- « *Structure d'organisation d'un document : configuration de ressources dans un espace de présentation* »
- « *Structure de contrôle : le contenu se déploie dans une dynamique de lecture imposée ou suggérée au lecteur* »
- « *Structure d'interaction : le contenu suscite une activité de la part du lecteur. Le support est un environnement qui offre et instrumente un espace pour la production d'informations par le lecteur dans le cadre d'une activité d'appropriation* ».

L'abstraction de ces différentes structures transforme alors le document en une interface très flexible, avec des coûts de développement très faibles. En comparaison, dans un contexte de base de données, « *l'accès au travers des formulaires de requêtes nécessite un important travail de développement et aboutit à des solutions rigides qui ne satisfont que des besoins bien précis et définis d'avance* » (Falquet 1999). Ainsi, le document numérique structuré, tel que nous l'avons abordé dans ce chapitre, peut être exploité comme une interface entre l'humain et la machine. On mettra notamment en avant que « *l'interface permet de manipuler les données, mais la vision document tend vers la définition d'une restitution plus adaptée à l'apprentissage et la compréhension* » (Zacklad et al. 1996).

## Chapitre 4 - Ingénierie des Interfaces Homme Machine

### 1 - Interaction Homme Machine

L'interaction homme machine est indissociable de la notion d'interface. L'interface dans une acception informatique peut être définie comme une « *jonction entre deux matériels ou logiciels leur permettant d'échanger des informations par l'adoption de règles communes physiques ou logiques* » (source : *Trésor de la langue française informatisé - <http://atilf.atilf.fr/tlf.htm>*). Elle peut également être définie comme « *logique intermédiaire entre deux traitements distincts, ou entre un traitement et un fichier* » (source : *Trésor de la langue française informatisé - <http://atilf.atilf.fr/tlf.htm>*). La première définition fait apparaître le fait que l'interface est un point d'ancrage entre des entités pouvant être physiques et logicielles. Ce point d'ancrage est défini pour permettre l'échange en s'appuyant sur des règles qui sont qualifiées de physiques ou logiques. L'interface peut en effet correspondre à la spécification d'éléments d'interconnexion, comme des prises dans une dimension physique, ou à la définition d'une logique implantée ou programmée, pouvant prendre la forme d'un protocole de communication réseau. Dans la dimension logique, on se rapproche du logiciel et de la manipulation et la communication de données abstraites stockées en mémoire. La seconde définition se concentre plus sur cette dimension logicielle et conceptuelle évoquant les notions de traitement et de fichier, qui correspondent en fait à la manipulation de données. A ce niveau, l'interface est implantée pour mettre en œuvre le concept d'abstraction inhérent à l'informatique. L'accès aux fonctionnalités de bas niveau de la machine se fait en traversant un empilement de couches d'abstractions accessibles chacune via des interfaces.

Au plus haut de cet empilement se trouve l'utilisateur humain. On aborde alors la problématique de l'interaction homme machine. Dans ce champ disciplinaire l'interface peut être définie comme « *le moyen de communication entre l'homme et le système* » *Coutaz et Ménadier cité dans (Zacklad et al. 1996)*. La notion de système est introduite ici. Elle représente une notion plus générale qui diminue la connotation matérielle et technologique de l'expression « Interface Homme Machine » pour élargir aux problématiques d'accès aux données contenues dans la machine. L'interface est ici vue comme un langage, les

mécanismes interactifs disponibles n'étant alors que des moyens de ce langage pour accéder, commander, communiquer.

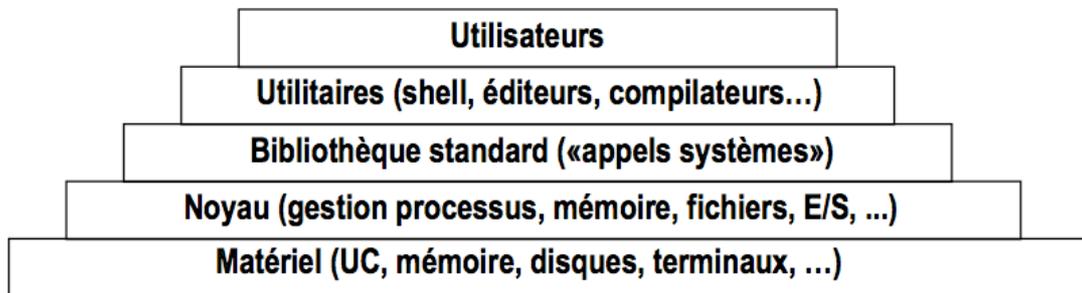


Figure 20 Illustration de l'empilement des couches d'abstraction dans un système d'exploitation

Le domaine de l'interaction homme machine est un champ disciplinaire très vaste que l'on peut toutefois dissocier selon qu'il se focalise sur les dimensions matérielle ou logicielle évoquées précédemment. Dans sa dimension matérielle, l'interaction homme machine s'intéresse à l'étude et au développement de périphériques permettant de communiquer avec tout type de machine. Il s'agit par exemple des périphériques de pointage ou de saisie, l'exemple caractéristique étant la *souris* présentée en 1963 et inventée par Doug Engelbert. Dans cet axe de recherche, c'est notamment l'efficacité des dispositifs produits au regard de la tâche à instrumenter qui guide les développements, l'orientation étant d'améliorer la performance de l'humain dans la réalisation de cette tâche.

L'interaction homme machine peut également être abordée selon une dimension plus logicielle et conceptuelle. En s'appuyant sur les couches d'abstractions existantes et déjà implantées, on s'écarte sensiblement de la dimension matérielle pour se rapprocher de la logique programmée. L'orientation est alors de se recentrer sur l'interaction homme machine, comme un langage, et de se concentrer sur la manière de communiquer les données abstraites contenues en mémoire à un utilisateur. Pour aborder ce champ d'étude, à l'expression interaction homme machine nous préférons *interaction homme information*.

En introduisant, l'interaction homme information, nous souhaitons aborder l'interaction au regard d'une vue linguistique plutôt que de se concentrer en premier lieu sur la tâche à réaliser. En effet, les concepteurs d'interface « *se sont surtout intéressés jusqu'à récemment, à augmenter leurs fonctionnalités et leurs performances* » (Kovacs, Gaunet et al. 2004). De

plus, la conception d'interface se focalise souvent sur l'activité de l'utilisateur appréhendée comme « *l'ensemble des moyens mis en oeuvre par l'utilisateur pour atteindre ses objectifs* » (Kovacs et al. 2004). Les interfaces produites constituent alors des environnements très cadrés, voire même contraignants, qui sont conçus et développés par rapport à des structures de données implantées sans prendre assez en compte la dimension utilisateur et sociale dans lesquelles les interfaces s'insèrent. Ce type d'approche fait partie des facteurs introduisant des difficultés de compréhension et d'appropriation du message que véhiculent les interfaces ou encore les fonctionnalités proposées.

En effet, dans une vue alternative, l'interface peut être abordée comme « la partie manipulable par l'usager qui complète un ensemble d'outils pour lui conférer le statut d'application » Nanard dans (Zacklad et al. 1996). C'est le point de vue de l'ergonomie qui apparaît ici. L'*outil* doit être considéré au sens de « *tout programme d'intérêt très général applicable à la résolution de problèmes divers* » (Nanard 1990) en concordance avec la notion d'outil introduite par Rabardel (Béguin et Rabardel 2000), c'est à dire un artefact conçu pour assister une tâche. Vient ensuite la notion d'*application* définie comme « *tout programme spécialisé dans le traitement d'un type de problème concret donné, l'homme intervenant explicitement dans ce traitement* » (Nanard 1990). Cette notion d'application prend en considération l'intervention humaine dans l'utilisation et la valorisation de l'outil proposé. En extension, cette notion d'application peut être rapprochée de la notion d'*instrument*, proposée par Rabardel qui doit être prise en compte dans le développement d'une interface. En effet, Rabardel introduit ce passage de l'outil à l'instrument lorsque l'outil intègre une activité métier. L'utilisateur va se l'approprier et surtout construire des schèmes d'utilisation de l'outil qui lui sont propres. Les schèmes d'utilisation correspondent à des modes d'utilisation de l'outil propres à chaque utilisateur, développés dans le contexte de l'activité. Ainsi, « *l'activité effective de l'outil est le plus souvent irréductiblement différente de la tâche initialement prévue. L'agent redéfinit la tâche et l'exécute d'une façon particulière* (Kovacs et al. 2004).

La problématique de l'interface ne se cantonne alors plus au développement d'une représentation graphique pour faire transiter des informations ainsi que les mécanismes permettant de les manipuler (Zacklad et al. 1996). L'interface s'insère dans une logique de système pas uniquement limitée au système informatique. La notion de système s'étend à l'activité métier et donc à la prise en compte d'une dimension humaine introduisant une *variabilité inter individuelle* (Kovacs et al. 2004) et une dimension sociale et culturelle. Ces facteurs rendent le problème moins établi du fait que l'interface doit être capable de supporter

le changement. Dans ce contexte, « *les méthodes classiques de conception pour l'utilisabilité fondée sur l'analyse et la modélisation des tâches n'accordent que peu de place à la gestion du changement et de son rôle dans l'organisation.* » (Bourguin et al. 2005). Cette gestion du changement met en avant l'importance de la flexibilité et de la *maléabilité* (Bourguin et al. 2005) des interfaces développées. Dans cette direction, la conception d'interface selon l'approche *modèle vue contrôleur* (MVC : model view controller (Krasmer et Pope 1988)) constitue une avancée introduisant une indépendance entre les données, leur traitement et leur visualisation :

- Le *modèle* traduit la couche applicative en se focalisant sur les données et leur bonne structuration. Il inclut également un ensemble de services, permettant de s'interconnecter avec d'autres systèmes par exemple.
- La *vue* a la possibilité de communiquer avec le modèle pour accéder aux données. Elle est en charge de calculer une présentation des données, c'est à dire produire une visualisation de données à l'intention des utilisateurs. Il s'agit de la couche interface utilisateur. À proprement parler, il s'agit de parler de « vues » au pluriel car l'indépendance de cette architecture offre une flexibilité, une maléabilité, permettant d'offrir un éventail de visualisations appropriées à chaque situation.
- Le *contrôleur* est en relation avec la vue et le modèle : il permet de répercuter sur le modèle les actions réalisées par l'utilisateur sur la vue. Il gère également au gré de l'interaction la succession des vues à présenter. Il intègre ainsi la dimension logique de présentation.

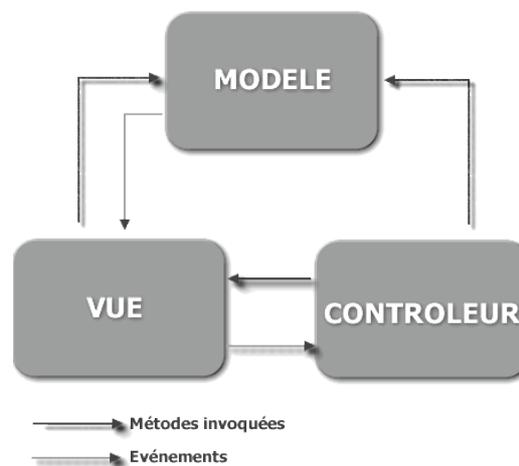


Figure 21 Architecture Modèle Vue Contrôleur (<http://www.labo-sun.com/resource-FR-articles-596-0-java-j2ee-introduction-a-j2ee.htm>)

Malgré l'apparente flexibilité de ce type d'architecture, une limite reste toujours que l'adaptation d'une interface n'est accessible et reste réservées aux concepteurs (Bourguin et al. 2005). De plus, l'approche MVC est majoritairement mise en oeuvre pour adapter la visualisation des données à leur manipulation et leur lecture, notamment dans des domaines où la tâche est bien balisée. Le fait que la tâche soit bien balisée permet de recenser avec précision les données ainsi que les opérations qui devront être disponibles. L'interface qui est alors construite et proposée découle, souvent des structures brutes de données mises en jeu ou d'une saisie procédurale : l'exemple caractéristique est le formulaire.

Ce type d'approche dans un contexte bien balisé appréhende toutefois l'interaction comme très codifiée. On considère alors trop souvent ces interfaces comme intuitives, faisant totalement disparaître les structures de données mises en jeu. On observe pourtant qu'un utilisateur non informaticien peut avoir d'importantes difficultés à comprendre le découpage inhérent aux contraintes de l'informaticien. On est alors confronté aux « *limites des approches classiques des IHM basées sur une psychologie de l'humain en tant que processeur d'information* » (Bourguin et al. 2005). Par exemple, lors du passage d'une commande via internet, on sera classiquement confronté à une succession de formulaires simples, chacun précisant les informations à fournir. Ceci constitue un domaine bien balisé et l'interface paraît acceptable du fait que le processus est commun et connu des utilisateurs. On s'attend en effet, après avoir validé son panier d'achat, à saisir les informations de livraison et de facturation, puis les informations de paiement. Dans un contexte beaucoup moins défini ou plus complexe, tel qu'un système mettant en jeu une modélisation complexe, la situation sera différente. Il s'agira de réussir à appréhender, par exemple, via une succession de formulaires, le processus et la sémantique. On identifie ici tout l'enjeu de faire appréhender à l'utilisateur non pas uniquement les structures de données mises en jeu mais bien la modélisation et le sens qui y est associé. En ce sens, plutôt que de se focaliser sur l'analyse de l'activité très liée à la manipulation des données, nous préférons nous intéresser à une vue plus linguistique de l'interface. Nous proposons donc dans la section suivante une étude de cette approche linguistique de l'interface qui va nous amener à considérer l'interaction homme machine comme une métacommunication.

## 2 - De l'interaction homme machine à la métacommunication

La vue linguistique de l'interaction homme information introduit une analogie entre interaction et dialogue entre individus. Cette analogie centre la problématique du

développement des interfaces sur la définition de langages. Les mécanismes interactifs proposés, ainsi que les modalités de visualisation de l'information mises en œuvre, ne constitueront que des éléments de ce langage permettant d'initier un dialogue entre l'homme et la machine

## 2.1. Dialogue homme machine

L'étude du dialogue homme machine renvoie aux recherches en intelligence artificielle et à l'apparition des premiers programmes interactifs. Les programmes interactifs se caractérisent par la possibilité d'interagir avec le programme durant son exécution se démarquant d'une conception « donnée – programme - résultat ». Un des premiers systèmes de dialogue est le système ELIZA qui représente un des systèmes de conversation les plus célèbres. Par la suite, des interfaces en langage naturel ont été développés pour accéder aux systèmes experts. Ils étaient basés sur l'utilisation d'ATN (Augmented Transition Network) Sémantique, la génération des phrases étant produite via des patrons de phrases. Un exemple de système extrêmement abouti à ce jour est le système Marlowe, qui correspond à un assistant virtuel d'analyse de discussions argumentées (Chateauraynaud 2003), exploité dans le domaine de la sociologie.

Le dialogue homme machine peut être abordé selon plusieurs orientations (Nicolle et Dizier 1999; Nicolle 2003) qui sont :

- Le dialogue entre humains et agents logiciels
- Le dialogue entre humains à travers les machines
- Le dialogue entre agents logiciels.

Nous ne présenterons pas une description exhaustive des recherches dans le domaine du dialogue homme machine. Il s'agit d'un domaine vaste, initié en intelligence artificielle et en linguistique, qui a été largement approfondi dans le cadre du développement de systèmes multi-agents par exemple. Notre positionnement sur l'interaction homme information nous conduit à nous intéresser au dialogue entre humains et agent logiciel même si nous verrons que nous aborderons la problématique du dialogue entre humains au travers des machines.

Initialement, l'étude du dialogue traitait essentiellement du langage oral, le langage écrit n'étant pas considéré comme du dialogue écrit en particulier par la psychologie. La tendance s'est inversée avec l'apparition d'outils de communication tel que le courriel ou la messagerie

instantanée (Chat), donnant lieu à de nombreux travaux en communication médiatisée par ordinateur (CMO) (Baker et al. 2001). La communication écrite peut donc désormais être appréhendée comme un dialogue.

Un dialogue est constitué d'un échange alterné entre au moins deux sujets (Nicolle 2003) : le locuteur (celui qui diffuse) et l'allocutaire (celui qui interprète). Le sens construit dans le dialogue dépend alors d'une co-construction entre les participants (Trognon et Brassac. 1992). L'étude du dialogue homme machine doit donc considérer les énoncés écrits de manière isolée, car ils constituent les éléments d'interaction pouvant être appréhendés comme des actes de langage (Searle. 1972; Searle et Vanderveken 1985). Il s'agit également de prendre en considération la structure globale du dialogue qui introduit différentes phases, comme la reconnaissance mutuelle, la proposition de sujet, l'acceptation du sujet ou encore la clôture du dialogue. Enfin, il est essentiel de considérer l'enchaînement conversationnelle. Cet enchaînement entre deux énoncés est, en effet, le facteur qui permet d'assurer une compréhension mutuelle entre les participants au dialogue.

La compréhension mutuelle est en relation avec les informations véhiculées entre les participants. Celle-ci dépend des théories du transfert « *qui traitent de la manière dont l'information est véhiculée entre les participants* » (Nicolle et Dizier 1999). Il est ainsi possible d'appréhender le transfert d'information selon 3 aspects :

- physique,
- logique ou social,
- sémiotique.

La *physique de transfert* correspond à la problématique de transmission d'information qui peut être considérée uniquement au niveau signal. Le cadre d'analyse est celui de la théorie de l'information basée sur le modèle de communication de Shannon et Weaver. Le modèle est constitué d'un émetteur et d'un récepteur qui communiquent au travers d'opérations de codage et décodage. L'orientation est ici de transmettre dans des conditions optimales un message à travers un moyen de communication donné. Le code est entièrement défini et doit être connu des participants au dialogue. Jakobson (1961) fait évoluer ce modèle de communication en introduisant la notion de référent. Le code n'est alors plus seulement appréhendé au niveau du signal, mais doit prendre en compte la dimension sociale et sémiotique. En effet, un élément du code, un symbole, a une signification dans un contexte,

un référentiel particulier, lié aux acteurs mis en jeu, ainsi que le contexte social dans lequel ils baignent.

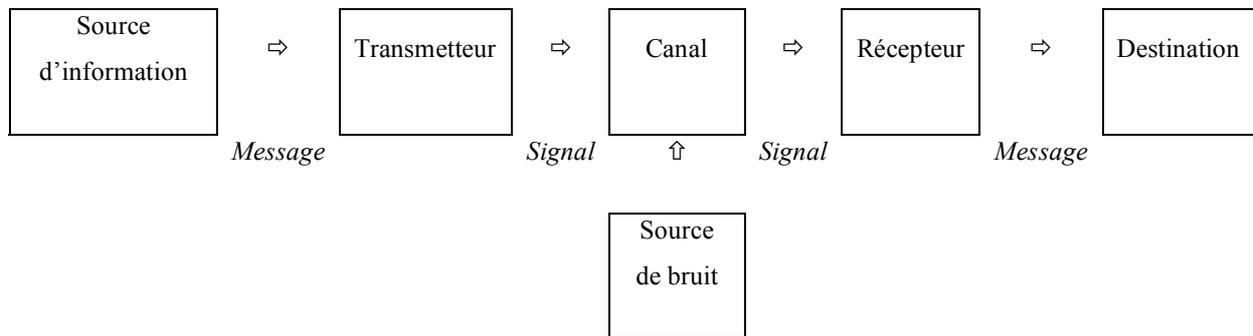


Figure 22 Modèle de communication de Shannon et Weaver (1948)

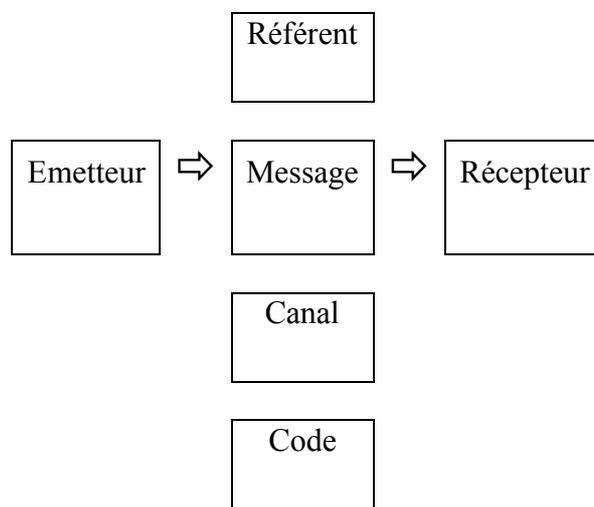


Figure 23 Modèle de communication de Jakobson (1961)

Le niveau physique du transfert permet ainsi d'identifier à la fois *les propriétés du medium* utilisé, ainsi que *les propriétés du signal* qui peut y être transmis. Ce signal qu'il soit auditif ou écrit constituera la représentation des informations à transmettre. Son interprétation nécessite *la connaissance d'un code*, qui constitue un métalangage définissant les règles d'organisation des symboles pour produire un énoncé signifiant.

La physique du transfert reste toutefois à un niveau bas d'analyse et appréhende la communication selon un schéma peu en accord avec la réalité. Boullier dans le domaine de l'ingénierie pédagogique évoque ainsi que « *certain schémas anciens ont la vie dure pour décrire la communication et celui de Shannon et Weaver où l'étudiant est alors traité comme un récipient qui se remplit, sans autre travail, des contenus qui lui sont déversés. Ce schéma de la communication technique continue de faire des ravages dans les représentations des processus éducatifs : les révisions systémiques (il y a du feed-back) ou à la mode de Jakobson*

(il y a codage et décodage) ne cherchent pas à remettre en cause les évidences selon lesquelles il serait possible d'identifier un émetteur et un récepteur, » (Boullier 2001). En effet, la communication ne peut être appréhendée comme unilatérale puisque communiquer c'est avant tout « agir sur l'autre » (Bateson et Winkin 1984). L'introduction de la *notion de référent* est toutefois intéressante dans le modèle de Jakobson puisqu'elle introduit déjà, pour partie, la dimension sémiotique que nous aborderons par la suite.

La *logique de transfert* renvoie à la dimension sociale du dialogue. Il s'agit de caractériser dans l'enchaînement conversationnel l'inversion des rôles de locuteur et d'allocutaire des participants. Le locuteur est en position de diffusion de l'information. Il produit des signaux avec un code qu'il a choisi ou sur lequel il s'est formé un consensus par exemple dans la phase initiale du dialogue. L'allocutaire se place en position d'interprétant. Il s'agit pour lui d'associer un sens aux signaux qu'il perçoit. L'allocutaire aura alors tendance à interpréter les signaux perçus selon ses spécificités. C'est à ce stade que nous pouvons introduire l'aspect sémiotique du transfert.

La *Sémiotique du transfert* est au cœur du processus de construction du sens entre les participants au dialogue. Elle renvoie aux problématiques de représentation, représentation du monde et à la construction de terrain commun. Lorsque deux individus dialoguent, chacun se construit une représentation de l'autre et de l'environnement dans lequel le dialogue a lieu : c'est ce que nous nommerons une *représentation du monde*. Elle est propre à chacun. Elle est influencée par l'origine et l'histoire de chaque individu. C'est à partir de là que se construisent les systèmes de référence sur lesquels se basent l'interprétation. L'ambiguïté du langage naturel, qui est polysémique, peut poser des difficultés dans la construction de ce terrain commun. Chaque acteur interprétera les énoncés selon son propre système de référence. Ceci explique les situations d'incompréhension, par exemple, entre les participants au dialogue qui vont permettre de faire évoluer les systèmes de référence de chacun pour développer un terrain commun entre les acteurs. C'est au sein de ce terrain commun que doit se placer la définition du code inhérent à la physique du transfert, ainsi que la définition et l'alternance des rôles de la logique du transfert. Le terrain est donc en évolution constante venant faire évoluer la limite d'un modèle de transmission basé sur un code déterministe, puisque celui-ci va se construire dynamiquement : cette vue de la communication peut être rapprochée du modèle de Riley (1959).

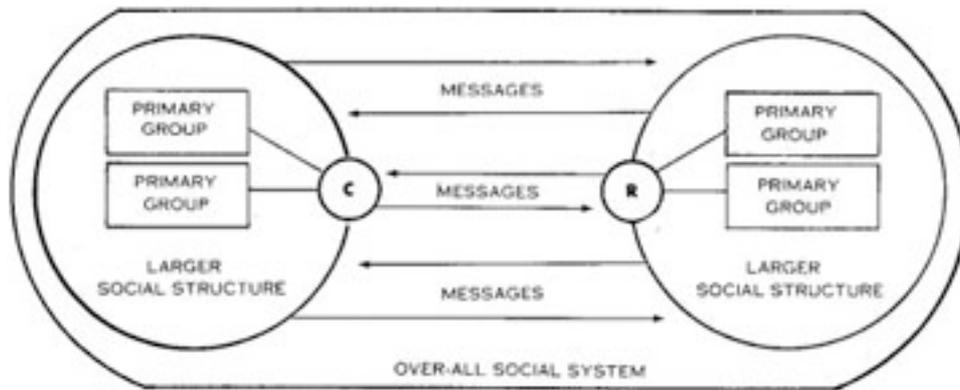


Figure 24 Le modèle de Riley (source :

<http://extension.missouri.edu/explore/comm/cm0109.htm> )

Dans le développement d'un dialogue homme machine, il s'agit d'attacher un soin particulier à la sémiotique du transfert même si le mode d'interaction au travers de l'écran, du clavier et de la souris ne permette pas une construction du terrain commun aussi aisée qu'entre deux individus. Cependant, homme et ordinateur possèdent leur propre représentation du monde, celle de la machine étant beaucoup plus figée, même si elle peut évoluer, mais au prix de mécanismes complexes, comme des systèmes d'apprentissage. Dans bien des cas, la représentation du monde de la machine ne peut évoluer que modestement, rendant difficile une réelle construction, par les deux parties en présence, d'un terrain commun. Il s'agit alors, d'une certaine manière, d'anticiper la construction de ce terrain commun.

Si bien évidemment l'ordinateur ne peut faire évoluer sa représentation du monde dans une large mesure, il sera par contre capable de fournir des informations qui précisent la signification des signaux qu'il transmet. Il s'agit, dans un premier temps, de rendre explicite au niveau de l'interface le système sémiotique inhérent à la visualisation proposée ainsi que les mécanismes interactifs disponibles. Il est également nécessaire de s'appuyer sur les codes culturels et de domaine des futurs utilisateurs. L'interface doit alors se baser sur leurs systèmes de référence dans l'idée de faciliter leur compréhension et leur utilisation. Dans la section suivante, nous traitons plus précisément le lien entre sémiotique et interface.

## 2.2. Sémiotique et IHM

*« the system's interface can be perceived as a communication generation artifact » (Oliveira Prates et al. 1997)*

L'interface est ici évoquée comme pouvant être perçues à l'image d'un artefact générateur de communication. Dans cette perspective, la dimension sémiotique de l'interface est un

---

fondement du message qui sera véhiculé. De la bonne compréhension de ce message dépend l'appropriation de l'interface ainsi que son utilisabilité<sup>4</sup> (Mallein et Tarozzi 2002).

L'interface constitue, en effet, un ensemble signes qui va mettre chaque utilisateur face à une *situation sémiotique* (Enjalbert 1996), une situation d'interprétation de ce qui lui est proposé. Ce type de situation sémiotique au regard de l'interface est évoqué par de (Souza 1996) comme : « *From the signs exchanged through Input/Output screens, users build a cascade of interpretants [Peirce'31] that eventually crystallize into the user's assigned meaning for the interface message --- the Application's Usability Model* ». Ainsi, l'ensemble des signaux véhiculés par l'interface, en entrée et en sortie, au travers de l'écran, peuvent être appréhendés comme des signes. Ces signaux tiennent tout aussi bien aux éléments qui constituent la forme de visualisation proposée qu'aux mécanismes interactifs implantés et disponibles.

Il est alors crucial que l'utilisateur interprète ces différents signes de la manière dont les concepteurs de l'interface les ont imaginés : ceci renvoie à la problématique de compréhension entre deux énoncés évoqués précédemment en dialogue. De plus, il est nécessaire que ce travail interprétatif ne soit pas trop lourd et complexe, au risque d'induire des charges cognitives importantes chez l'utilisateur (Souza 1996). Tout ceci souligne l'importance de ne pas concevoir et développer l'interface uniquement au regard de la tâche et de la manipulation des données mais également de véhiculer au travers de son utilisation le niveau métalinguistique qui caractérise la sémantique des signes dont elle est constituée (Souza 1996).

Ainsi, « *il est de la responsabilité de chaque concepteur d'application de se soucier de l'homogénéité du système sémiotique qu'il utilise* » (Zacklad et al. 1996). Cette homogénéité est largement pratiquée (cf. Figure 25) dans le développement des interfaces WIMP (« Windows, Icons, Menus and Pointing device ») qui introduisent par exemple une standardisation des icône facilitant leur compréhension : les icône répondent alors à un « *code sémiotique standardisé* » (Zacklad et al. 1996) . Cette notion est également pratiquée au niveau des mécanismes interactifs, tel que raccourcis clavier répondant aux mêmes combinaisons de touches d'une application à une autre, d'un système d'exploitation à un autre.

---

<sup>4</sup> L'utilisabilité est à différencier de l'utilité qui correspond au respect de l'atteinte du but. Les critères d'utilisabilité sont la facilité d'apprentissage et d'utilisation, l'efficacité d'utilisation, la facilité de mémorisation, l'utilisation sans erreurs, la documentation, la satisfaction, Kovacs, B., F. Gaunet et X. Briffault (2004). Les techniques d'analyse de l'activité pour l'IHM, Hermès Science.

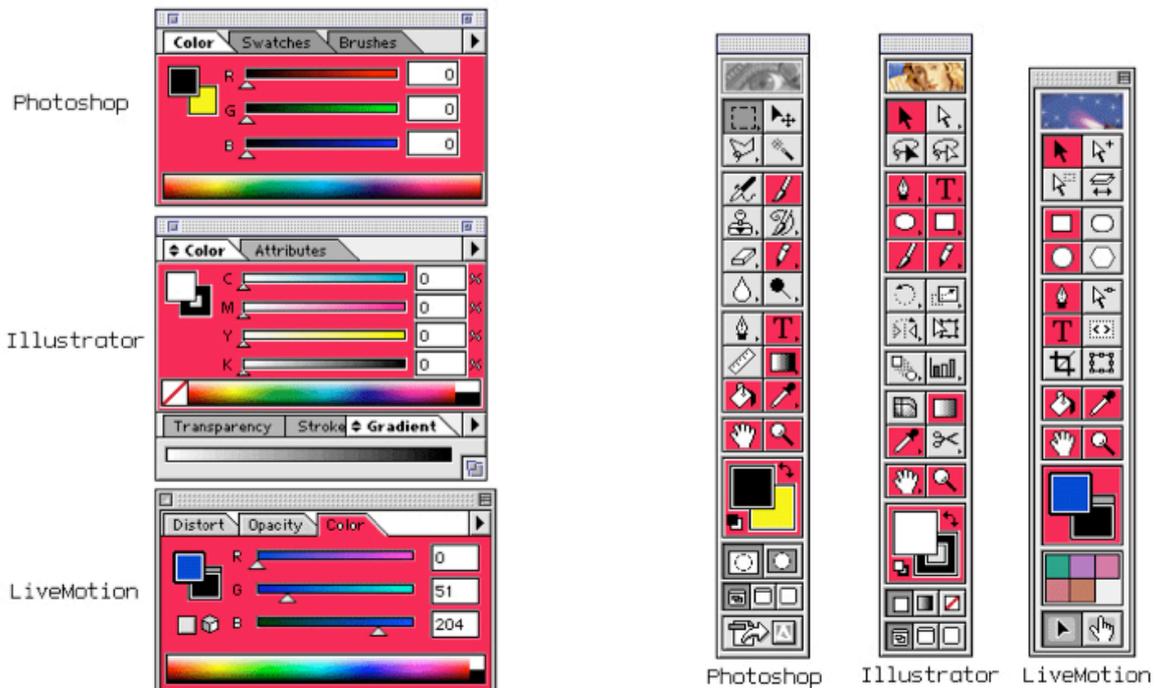


Figure 25 La standardisation des icônes entre différentes applications Adobe (source :  
(Beaudoux 2004) )

Si cette standardisation est pratiquée dans de nombreuses applications largement déployées, dans le cadre de visualisation d'information et des mécanismes interactifs communs et bien balisés (copier = ctrl + c, coller = ctrl + v), elle peut être plus délicate dans des applications métier. La dimension métier introduit nécessairement à la fois des approches et pratiques spécifiques sur le plan de la manipulation et du traitement de l'information, mais également des vocabulaires et formes de représentation particulières. Il s'agit alors dans la conception et le développement des interfaces de s'appuyer sur des « *structures sémiotiques connues* » (Zacklad et al. 1996) pour faciliter l'appropriation de l'interface par les futurs utilisateurs.

Ce point appuie le fait que « *la définition du système sémiotique nécessite une analyse indépendante des moyens informatiques* » (Zacklad et al. 1996). Il semble ainsi « *nécessaire de définir des concepts sémiotiques indépendants des objets informatiques qui permettront à l'analyste de définir le système qui lui semble le plus adapté à ses besoins* » (Zacklad et al. 1996). L'indépendance qui est préconisée ici doit permettre de mieux identifier et qualifier les éléments constituant le message que va véhiculer l'interface. A partir de là, la mise en relation des concepts sémiotiques avec les objets informatiques constituera le niveau métalinguistique de l'interface. L'approche proposée dans (Zacklad et al. 1996) est alors de « *rapprocher un* »

---

*modèle conceptuel de résolution de problème " d'un " modèle du système sémiologique adapté au problème à résoudre " pour arriver à un modèle conceptuel synoptique ». Les auteurs s'intéressent ici au développement de représentation synoptique pour la supervision de réseaux de télécommunication. Cette approche de la conception d'interface peut être rapprochée de l'approche MVC, mais se centre plus sur une vue linguistique. Ainsi, le niveau métalinguistique défini constitue une réelle légende de l'interface. Le lien entre les données (objets informatiques) et leur représentation (système sémiologique) n'est plus noyé dans l'implémentation, ce qui auparavant le rendait tacite pour des non initiés. Il est alors possible d'imaginer que la malléabilité de l'interface ne soit plus réduite aux concepteurs en rendant disponible et modifiable cette couche métalinguistique (modèle conceptuel synoptique). Cette idée reste toutefois encore difficile à réellement mettre en œuvre avec les environnements de développement actuels. On reste souvent tributaire des concepteurs d'interface. Toutefois, il est possible de leur proposer certaines recommandations dans le cadre d'une ingénierie sémiotique. En voici deux exemple issues de (Souza 1996) :*

- « *Designers should avoid using invented interface signs; »*
- « *Designers should select interface token sign systems from culturally established type systems of communication; »*

La première recommandation évoque l'importance d'éviter à chaque nouvelle interface de mettre en œuvre des signes différents pour des fonctions identiques. Il s'agit de s'attacher une continuité dans le développement des interfaces afin d'en faciliter l'appropriation. Ceci renvoie à l'importance de conserver une homogénéité dans les systèmes sémiotiques mis en œuvre que nous avons évoqués précédemment.

La seconde recommandation spécifie l'importance de s'appuyer sur des structures sémiotiques connues. Il s'agit de prendre en compte la dimension culturelle dans le développement d'interface et donc de s'appuyer sur les langages, les formes de représentation propres aux métiers, aux domaines des utilisateurs lorsque cela est possible et lorsque de tels codes existent.

Ces recommandations sont empreintes de la philosophie sémiotique évoquant largement la problématique du signe et donc une vue linguistique de l'interface. Ainsi, en tenant compte des acteurs qui gravitent autour de l'interface, s'agissant tout aussi bien de ces concepteurs que de ces utilisateurs, elle peut être présentée comme « *a one way message sent from the*

---

*designer to the user, conveyed through a computer media. The system's interface is a message that can also exchange message with the user ; hence it is a metacommunication artifact » (Oliveira Prates et al. 1997). L'interface n'est plus ici présentée comme un dispositif permettant de manipuler les données, mais comme un message entre les concepteurs et les utilisateurs. Ainsi, chaque élément constituant l'interface véhicule une sémantique et une intention définie par les concepteurs à destination des utilisateurs. Ces utilisateurs doivent alors être en mesure d'interpréter correctement le message proposé : l'interface est alors vue comme une métacommunication.*

### 2.3. L'interaction homme machine vue comme une meta communication

Selon Paul Dourish « *human computer interaction can be thought of as a form of a mediated communication between the end user and the system designer » (Dourish 1999). Ce positionnement qui s'ancre dans le domaine du CSCW et du Social Computing est tout à fait en accord avec la vue linguistique de l'interaction homme machine mettant en avant l'enjeu crucial de bien définir le niveau sémiotique de l'interface.*

Cette manière d'appréhender l'interaction homme machine est également partagée dans le domaine de l'ingénierie éducative où l'on peut voir apparaître la *notion d'intermédiation* dans le développement des systèmes interactifs (Bourguin et al. 2005). L'intermédiation met l'accent sur le fait qu'une séparation entre les utilisateurs et les concepteurs n'est pas une bonne solution. Il s'agit donc d'intégrer dans le processus de conception et de déploiement des interfaces « *une intermédiation (un passage plus qu'une séparation) entre les concepteurs et les usagers » (Bourguin et al. 2005). Dans la conception, l'intermédiation se positionne « entre les besoins, compétences, désirs... des usagers et les contraintes, technologiques ou autre, les compétences et les désirs des concepteurs. Le système est une inscription dans la technologie qui traduit l'équilibre local, temporaire, négocié, fragile entre les parties ; » (Bourguin et al. 2005). Dans le déploiement, l'intermédiation se place « entre les deux parties, usagers et concepteurs, pour en assurer la maintenance et l'évolution dans la durée. » (Bourguin et al. 2005).*

Cette notion d'intermédiation au niveau de la conception et du déploiement met l'accent sur un aspect particulier du développement d'interface par rapport au développement des parties de systèmes informatiques plus centrées sur la manipulation de données et pas en relation avec l'utilisateur. Dans ce dernier cas, un exemple pouvant être le développement d'une base de données, il s'agit pour le concepteur de recenser les données à gérer dans le problème qu'il

traite pour ensuite concevoir un modèle relationnel, qu'il implémentera au sein d'un système de gestion de bases de données. Cette tâche peut introduire une séparation assez forte entre les concepteurs et les utilisateurs. Par contre, dans le cas de l'interface homme machine constitue la partie visible et manipulable d'un système informatique, elle constitue l'outil que l'utilisateur transformera en *instrument* au sens de Rabardel en développant ses propres stratégies d'utilisation (notion de Schèmes d'utilisation) et de mise en œuvre (Béguin et Rabardel 2000). Il est alors admis que l'outil va transformer la tâche rendant la définition des IHM très indéterministe. De là, les démarches de conception, trop prédictives et centrées sur l'analyse de la tâche pâtissent du fait de vouloir définir a priori une activité qui va se modifier. Ceci met en avant l'importance de travailler sur des cycles courts de conception qui, en fait, remettent sensiblement en cause la dichotomie entre conception et déploiement. Des approches telles que l'ingénierie concourante, la conception participative (O'Neill 2001) ou encore l'*extreme programming* (Casabianca 2005), centrées sur une implication et collaboration des utilisateurs avec les concepteurs, semblent alors beaucoup plus adaptées.

Dans cette ordre d'idée, Kaptelinin évoque : « *Perhaps a better understanding of differences between users and designers, between design for use and design in use, may help develop more practical versions of participatory design* » (Kaptelinin 2003). Cette citation introduit la variabilité de point de vue au regard de l'interface à produire. Le concepteur est majoritairement positionné sur une conception et un développement largement influencés par les structures de données mise en œuvre ainsi que la tâche qu'il a pu identifier : c'est le *design for use* guidé par l'*utilité*, l'interface devant permettre l'atteinte de buts identifiés (Kovacs et al. 2004). L'utilisateur, lui, est plus sensible à l'*utilisabilité* et donc au *design in use*. Il s'agit pour lui que l'interface ne réponde pas seulement à des buts, mais introduise également une certaine ergonomie, un certain confort dans son utilisation. De plus, il s'agit qu'elle s'intègre bien à son activité métier, sa pratique. C'est bien évidemment un compromis entre les deux points de vue, qu'il faut réussir à trouver, qui passe nécessairement par une compréhension entre concepteurs et utilisateurs. Les concepteurs doivent mieux comprendre l'activité des utilisateurs, alors que ces derniers doivent faire l'effort d'appréhender les contraintes des concepteurs, par exemple conceptuelles et techniques (ex : unicité, repérage, typage et cohérence en base de données) afin que leurs informations soient correctement gérées.

Cette compréhension évoquée dans le processus de conception doit également être prise en compte lorsque l'interface est utilisée, nous renvoyant à l'importance de la vue linguistique de cette interface. En effet, « *il est clair que dans la plupart des approches ou méthodes de*

développement classiques des IHM, l'utilisateur est laissé livré à lui-même pour tout ce qui suit la phase de tests et de déploiement (Wild, Macredie, 2000). » (Bourguin et al. 2005). Les interfaces sont alors plus ou moins acceptées, plus ou moins comprises selon qu'elles sont faciles à comprendre, ou encore selon qu'elles sont abordables. La notion d'abordance (Gibson 1977; Gibson 1979) est issue de la psychologie de la perception. Ainsi, « Le concept d'abordance a été inventé par le psychologue de la perception Gibson [1,2] pour désigner les propriétés actionnables entre le monde et un individu (personne ou animal) » (Norman 1999). Cette notion d'abordance est en fait issue de théorie écologique faisant référence aux animaux qui s'adaptent à leur milieu. Dans la conception d'un artefact ou d'une interface, une abordance peut être liée au fait que « l'apparence du dispositif pouvait fournir les indices critiques requis pour une utilisation appropriée » (Norman 1999). Par exemple, ceci se matérialisera par l'intégration dans un dispositif technique, tel qu'une poignée, de rainures ou d'encoches matérialisant la position des doigts. Il semble que « le concept d'abordance a été adopté par la communauté des concepteurs, et tout particulièrement dans le domaine de la conception graphique et industrielle » (Norman 1999). Mais force est de constater que « le designer se préoccupe plus que les actions que l'utilisateur perçoit soient possibles plutôt que réelles » (Norman 1999).

En fait, dans le domaine de la visualisation sur écran, mettre en place des abordances réelles et physique, comme dans l'exemple de la poignée, se révèle difficile en dehors de dispositif interactifs matériels en complément de la visualisation. Il s'agira de se limiter aux abordances perçues, renvoyant à la définition et la mise en place de contraintes et de conventions. Les conventions culturelles jouent notamment un rôle majeur au niveau de l'écran (Norman 1999). La mise en place de conventions, surtout dans la forme de représentation des informations, nous renvoie à la fois à la dimension sémiotique de l'IHM ainsi qu'à l'importance de les considérer comme une forme de communication médiatisée véhiculant un message entre concepteurs et utilisateurs. Ainsi, « User Interfaces are metacommunications artifacts. They are designed to convey a message from system designer to system user whose meaning is the answer to two fundamental questions: (1) "What kinds of problems is this application prepared to solve?" and (2) "How can these problems be solved?". (Oliveira Prates et al. 1997). Le message que l'interface doit permettre de percevoir, tel qu'il est présenté ici, tient à deux aspects :

- A quel type de problème cette interface est elle adaptée ?

- Comment le problème peut-il être résolu ?

Ces deux questions renvoient à un contexte où l'interface permet d'accompagner la résolution de problème. Dans un cadre plus élargi, notamment des interfaces de diffusion et de communication d'informations, il s'agira que le système sémiotique inhérent à l'interface soient aisément identifiable. Dans cette direction, « *il est donc nécessaire que le système interactif initialement développé puisse servir lui-même comme un pont entre les deux parties au travers de l'artefact lui-même. Cela implique que [...] le SIE doit contenir sa propre représentation (un méta-modèle)* » (Bourguin et al. 2005). Ce passage évoque la notion de « méta modèle » du système interactif (SIE : Système Interactif Evolutif). Le méta modèle doit constituer un contrat de lecture sur l'interface, le langage qu'elle définit, permettant de la percevoir de manière monosémique. Sa mise à disposition auprès de l'utilisateur doit accompagner l'utilisateur dans l'exploitation de cette interface. On pourra alors considérer que ce méta modèle est appréhendable naturellement, si, par exemple, l'interface s'appuie sur une codification ou des conventions acceptées et connues par les utilisateurs. En dehors de cette situation, il s'agira de le préciser sous la forme d'une légende ou d'une documentation disponible *in vivo*.

### 3 - Interface homme information

Nous avons évoqué au début de ce chapitre la notion d'interaction homme informations qui se focalise sur la mise en relation entre l'utilisateur et des données abstraites contenues en mémoire. En considérant l'interaction homme machine comme une métacommunication, nous proposons de dissocier deux types d'interface homme information :

- Les interfaces de lecture
- Les interfaces auteur

Les *interfaces de lecture* se focalisent sur la diffusion du message entre les concepteurs et les utilisateurs de l'interface. Dans la conception de ces interfaces, il s'agit de s'attacher aux modalités de communication des informations, que cela soit dans le choix du médium, du signe et de la forme. L'interface est adaptée, et souvent spécifique, au message et n'offre qu'une malléabilité limitée au travers de comportements dynamiques et interactifs. La modification ou l'adaptation du message transmis doit passer par les concepteurs.

Les *interfaces auteur* se différencient par le fait qu'elles vont permettre la communication d'un contenu, mais également l'édition du contenu. Il pourra s'agir ou d'interfaces permettant

---

de construire un message à partir de moyens de mis en forme de texte ou de graphique. Il pourra également s'agir d'une interface de diffusion de contenu, comme évoqué précédemment, mais permettant une contextualisation du contenu par l'utilisateur. Dans ce dernier cas, « *le contenu [véhiculé par l'interface] cesse d'être un objet indépendant du récepteur et immuable, pour devenir une matière à recomposer, comme le sont les cours en salles que chacun doit suivre, synthétiser et revoir selon sa compréhension et ses priorités* »(Dufresne 1997). Ainsi, Dufresne précise « *que l'environnement doit être malléable car il ne peut être défini a priori. Les interfaces pourraient ainsi être conçues comme des outils supportant directement l'activité et non plus uniquement comme des intermédiaires entre des outils, des services et des utilisateurs* » (Dufresne 2001).

Ce positionnement peut amener à considérer les interfaces comme des documents et à les construire comme tels. L'interface comme nous l'appréhendons constitue un medium de communication véhiculant un message (un signe) entre concepteur et utilisateur avec une présentation (une forme) particulière. En ce sens, l'interface peut être rapprochée de la notion de document réunissant ces trois dimensions. Elle se différencie en général du fait que les données en dehors de la perception sont dissociées de la présentation, contrairement à un document dont les données sont intégrées à la présentation, ce qui en fait un objet persistant (Beaudoux 2004) et autoportant. Le document conserve son état et doit être éditable pour une grande part sans être en relation avec un système. Ainsi, Beaudoux (Beaudoux 2004) affirme qu'il est « *inutile de posséder une application spécifique pour le lire* ». Ce dernier point est difficile à atteindre dans son intégralité puisque l'on est souvent confronté à des difficultés de compatibilité entre format dans le monde numérique.

La problématique de construction du message n'est pas simple si l'on tient compte en particulier de la dimension sémiotique de l'interface que nous avons discuté. Il s'agit de bien prendre en compte ces spécificités et de ne pas uniquement se limiter à une analyse des besoins où l'utilisateur se positionne essentiellement en source d'information : l'utilisateur doit se révéler une source d'inspiration dans la conception de l'interface et la définition des modalités du message qu'elle véhicule.

Le message véhiculé par l'interface est essentiellement lié à la perception directe et sémantique de la visualisation proposée à l'utilisateur. La *perception directe* renvoie à la forme de visualisation d'information : la présentation (Beaudoux 2004). La *perception sémantique* correspond « *aux données qui sont interprétées par l'utilisateur via la*

*présentation* » (Beaudoux 2004). C'est toute l'importance de la forme de visualisation, que nous allons aborder dans le chapitre suivant.

## Chapitre 5 - Ingénierie de la forme

La forme est liée à la perception directe qu'un utilisateur, un lecteur, va avoir de ce qui lui est proposé sur le support choisi. C'est par le biais d'une forme particulière, donnant une matérialité à une inscription, que la sémantique va être médiatisée. La forme correspond alors tout aussi bien à la mise en forme matérielle de texte, qu'à la définition d'une approche graphique de représentation de données abstraites. Nous proposons de nous concentrer ici sur la modalité graphique. L'idée n'est pas tant de s'intéresser à des langages déjà spécifiés de représentation, tels que le code de la route ou encore la visualisation de connaissances médicales (Lamy 2006). Nous préférons aborder la définition d'une forme de représentation graphique en nous intéressant tout d'abord à la visualisation d'informations, puis à la sémiologie graphique.

### 1 - Visualisation d'informations

La visualisation d'informations renvoie à un champ de recherche très vaste et qui prend un caractère crucial dans les problématiques d'accès et d'analyse d'informations. L'enjeu est de fournir une forme de représentation des informations permettant de les valoriser au mieux. Une visualisation d'informations constitue une représentation externe (Cox et Brna 1995) qui doit réduire la surcharge cognitive en mettant à contribution les capacités visuelles humaines. En effet, l'œil constitue un des dispositifs d'analyse de données les plus efficaces qui soient.

Ainsi, le domaine de la visualisation doit être une réponse aux données abstraites qui ne possèdent pas de représentations physiques directes. En effet, si dans de nombreux champs disciplinaires, la démarche expérimentale permettait de visualiser directement un phénomène (une expérience dans un tube à essais), la technologie et l'informatisation introduisent une information numérique qui constitue un apport certain, mais devant faire l'objet d'une traduction sous une forme particulière pour pouvoir être visualisée.

Ainsi, une visualisation, présentée précédemment comme une représentation externe, doit constituer un artefact cognitif. « *a cognitive artefact is an artificial device designed to maintain, display, or operate upon information in order to serve a representational function* » (Norman 1991). Une visualisation d'information constitue effectivement un artefact cognitif au sens où celle-ci permet de représenter de manière artificielle des données abstraites. Un

artefact cognitif est également présenté comme devant améliorer les capacités cognitives humaines. La notion d'amélioration ne doit pas être vue comme portant sur les capacités individuelles l'artefact cognitif ne faisant que modifier la tâche, la rendant plus simple ou plus adaptée aux capacités humaines.

La visualisation d'informations doit donc constituer un outil permettant le traitement de l'information, à l'image de la graphique que nous introduirons par la suite : c'est le « *using vision to think* » (Stuart K. Card et al. 1999). En ce sens, elle doit accompagner la découverte et la création d'idées (Morand 2000), favoriser la compréhension en introduisant l'explication et appuyer la décision (Tufté 1997). Par exemple, « *une contrainte non explicitée peut être découverte dans l'observation du dessin (reconnaissance d'une symétrie par exemple)* » (Morand 2000). Ce sont les capacités visuelles de l'humain qui sont mises à contribution mais cela suppose de produire un bon dessin. Ainsi, « *le support de représentation peut alors représenter un espace de recherche dans la résolution de problèmes* » (Morand 2000). Edward R. Tufté évoque « *the consequences resulted directly from the quality of methods used in displaying and assessing quantitative evidence* » (Tufté 1997). Les conséquences évoquées ici sont celles d'une prise de décision. L'accent est ici mis sur le caractère crucial de la visualisation qui prend, dès lors, une responsabilité énorme dans la prise de décision.

### 1.1. Classification des techniques de visualisation

Les techniques de visualisation développées sont nombreuses (Beaudoin-Lafon et Hascoët 2001) et le retour d'expérience où l'évaluation fait encore cruellement défaut. Il existe plusieurs initiatives comme des évaluations sur le plan cognitif (Le Blanc et al. 2003) ou encore des travaux plus liés à l'ergonomie de surface, qui ne se centrent pas nécessairement sur la problématique de la visualisation, mais plutôt des interfaces homme machine en général (Bastien 1991).

Nous proposons d'introduire une initiative de typologie des techniques de visualisation. Il s'agit en fait de construire une table classifiant les techniques, à l'image de la table de classification périodique des éléments chimiques de Mendeleev. Cette table, dont on peut observer une illustration sur la Figure 30, introduite dans (Lengler et Eppler 2007), propose six catégories de méthodes :

- Data visualisation,
- Information visualisation,

- Concept visualisation,
- Metaphor visualisation,
- Strategy visualisation,
- Compound visualisation.

La catégorie *Data visualisation* inclut les méthodes de visualisation de données quantitatives, comme les graphiques sous forme d'histogrammes, de secteurs, de courbes, etc... Les tableaux de données entrent également dans cette catégorie. Dans ce type de méthode, les dimensions du plan sont associées à un type d'information.

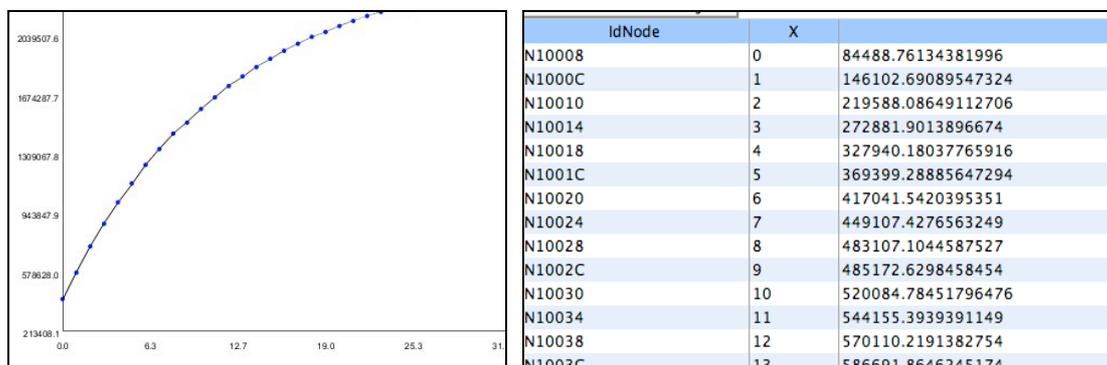


Figure 26 Courbe et tableau comme exemple de Data visualization

L'*Information visualisation* renvoie essentiellement à la visualisation interactive d'information. L'ambition est bien d'amplifier la cognition à travers une transformation des données sous une représentation visuelle pouvant être manipulée et transformée par l'utilisateur.

La catégorie *Concept visualisation* correspond aux méthodes permettant une visualisation d'information qualitatives majoritairement sous la forme de représentation graphique à deux dimensions. La visualisation de graphes (Mugnier et Chein 1995; Karouach et Dousset 2003) trouve ici toute sa place (cf. Figure 27).

La *Metaphor visualisation* introduit la notion de métaphore, au sens où la représentation de l'information va s'appuyer sur une image, une symbolique, connue et acceptée. L'objectif est de s'appuyer à la fois sur l'organisation et la structuration graphique inhérente à l'image connue, sur laquelle on s'appuie pour faire transiter la sémantique des informations que l'on veut représenter.

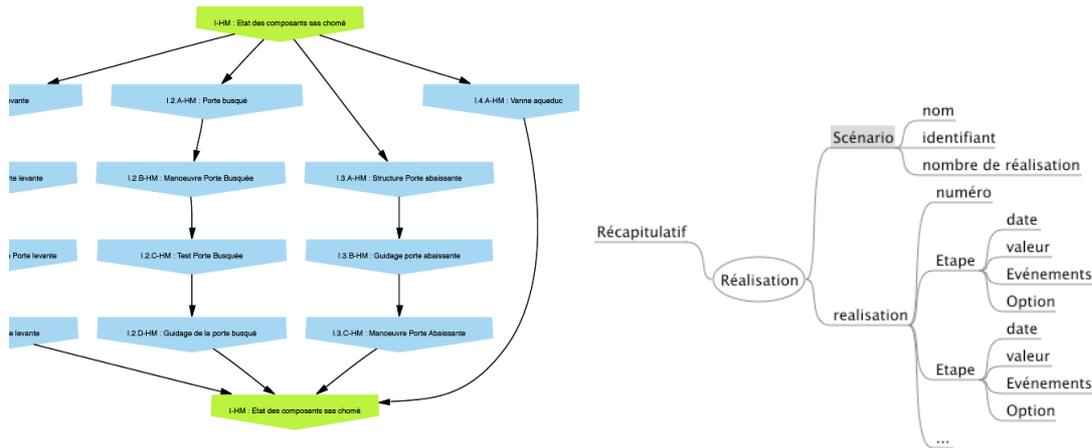


Figure 27 Exemple de visualisation de graphe et d'une carte conceptuelle

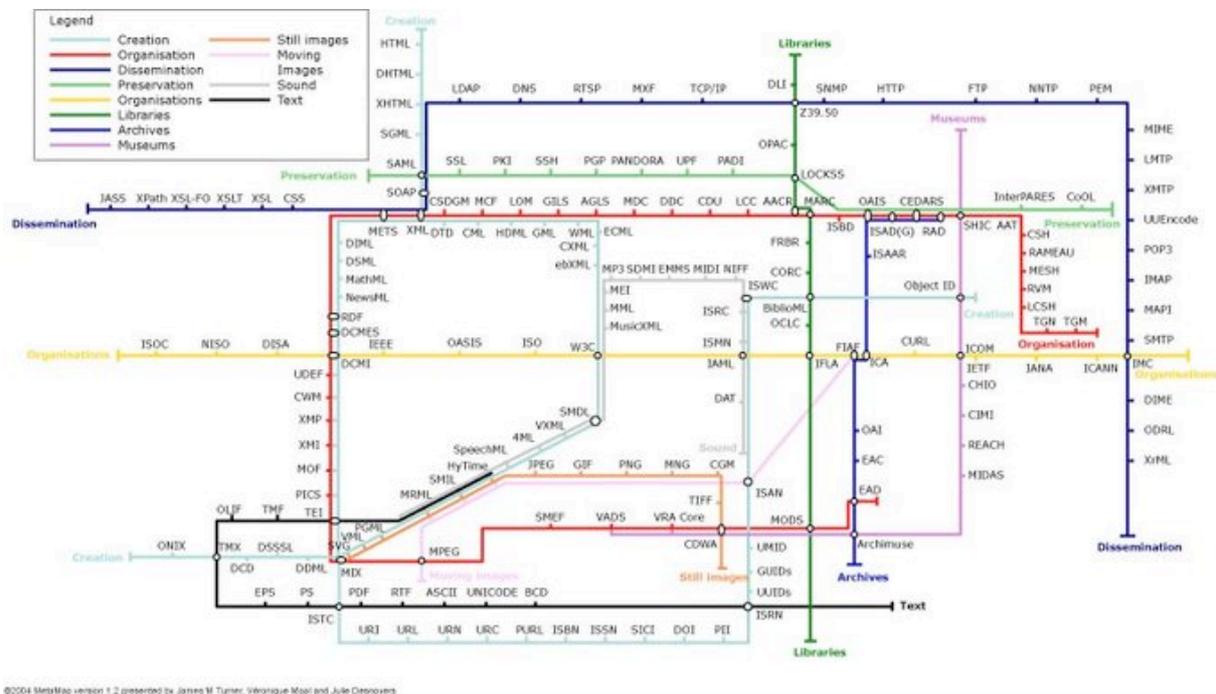


Figure 28 MétroMéta : un outil pédagogique de présentation de normes sous une forme graphique : une carte de métro

(<http://www.mapageweb.umontreal.ca/turner/meta/francais/index.html>)

La *Strategy visualization* est une catégorie qui renvoie aux méthodes de visualisation exploitées en particulier dans le domaine du management. Ces modes de visualisation exploitent la représentation visuelle pour améliorer l'analyse, le développement, la formulation, la communication et l'implémentation de stratégies d'organisation.

Enfin, la catégorie *Compound visualization* fait référence à des méthodes de visualisation intégrant différents types de représentations graphiques (ie. Différents type de méthodes de visualisation) dans la même image.

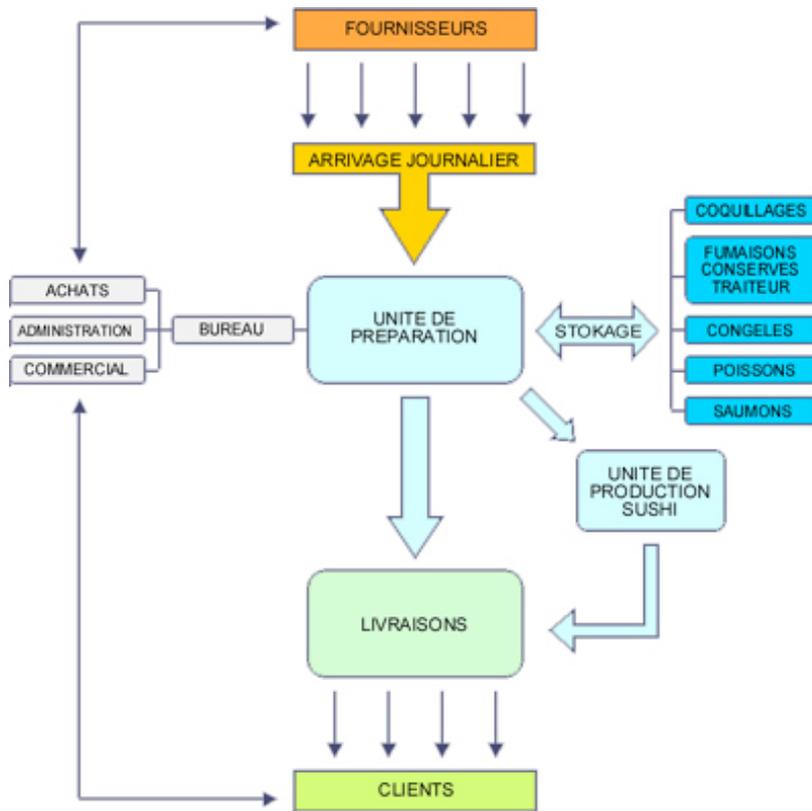


Figure 29 (source : <http://www.lucas.ch/>)

Les méthodes de visualisation ne sont pas uniquement associées dans les catégories que nous venons de présenter. La table propose également un typage des méthodes selon trois autres dimensions qui sont : le type de données à représenter, le type de tâche et d'interaction au regard des données et le type de processus cognitif que l'on souhaite stimuler. Nous proposons maintenant d'aborder la dimension interaction qui ne peut être dissociée de la visualisation.



## 1.2. Visualisation et interaction

Visualisation et interaction sont intrinsèquement liées. En effet, l'utilisateur lorsqu'il va être confronté à une visualisation va interagir avec elle au travers de la perception visuelle. La perception sera tout d'abord globale puis il s'attardera à analyser les composants de l'image : la perception sera alors plus locale.

Ainsi, lors de la conception d'une visualisation interactive, il s'agira de prendre en considération le Visual Information-Seeking Mantra (*Shneiderman 1996*) : « *overview first, zoom and filter, then detail on demand* ». Ce principe est fondateur dans la conception d'un système de visualisation interactif de données et a pris toute son importance et trouvé sa pertinence avec le succès des interfaces à manipulation directe.

Ainsi, l'*Overview* (vue d'ensemble) correspond au fait d'obtenir une vue d'ensemble des données visualisées. Il s'agit là de la capacité première d'une visualisation ou d'une représentation graphique qui offre une vue globale. La représentation graphique est toujours perçue au premier abord de manière globale. C'est le principe même de la théorie de la Gestalt (*Gurwitsch 1935*). Ainsi, une visualisation peut être appréhendée comme une forme ou Gestalt, c'est à dire « *un ensemble structuré et doté de sens* » (*Lamy 2006*) : cette ensemble vaut plus que la somme de ses parties expliquant la perception globale.

Le *Zoom* doit permettre d'obtenir une vue précise d'un élément composant la visualisation. Il offre la possibilité à un utilisateur de définir son centre d'intérêt pour en obtenir une vue plus précise. Il s'agit de « *concilier détail et vue globale* » (*Beaudoin-Lafon et Hascoët 2001*). Il est possible d'identifier différentes formes de zooms. Nous les classerons selon qu'ils correspondent à un *zoom infini* ou un *zoom sémantique* (*Beaudoin-Lafon et Hascoët 2001*).

Le *zoom infini* correspond à une opération géométrique sur une image bitmap (image raster) permettant un changement d'échelle sans pour autant transformer son contenu. Ainsi, un zoom avant révèle les détails alors qu'un zoom arrière offre une vue plus contextuelle. Le zoom infini introduit toutefois une rupture lors du passage entre vue globale et détail. L'introduction de technique de déformation, telle que le fisheye, permet d'intégrer une certaine continuité dans la perception du détail au sein même d'une vue globale.

Le *zoom sémantique* est une opération plus complexe qui va à la fois introduire une opération géométrique, mais également une opération logique sur les données. Lors d'une phase

d'interaction avec ce type de zoom la représentation des données va changer en fonction du niveau de détail. L'outil *zoomit* (S. Pook et al. 2000) implémente ce type d'approche transformant le zoom en un outil de navigation dans une structure de données. Point important, le zoom n'offre plus une réelle continuité, puisque, d'un pas de zoom à un autre, la visualisation peut changer. Il s'agit donc que la dimension logique du zoom soit bien explicitée afin de ne pas désorienter l'utilisateur.

Le *Filter* (le filtrage) renvoie au fait d'offrir la capacité de moduler les éléments visualisés. Cela doit permettre de faire apparaître ou disparaître tout ou partie des données qui ne sont pas jugées pertinentes par l'utilisateur. En effet, le choix des modalités de visualisation de données constitue un exercice difficile au sens où il est complexe d'espérer anticiper sur les besoins de l'utilisateur. Cette difficulté est très nette lorsque la visualisation doit être un support au traitement de l'information. En effet, elle risque de remettre en question les savoirs de l'utilisateur et celui-ci souhaitera certainement la modifier en fonction de sa réflexion. Ceci met l'accent sur l'importance d'intégrer des mécanismes de *filtrage dynamique* (Beaudoin-Lafon et Hascoët 2001) permettant d'adapter la visualisation en contexte.

Enfin, le *Details-on-demand* doit permettre de sélectionner un ou plusieurs éléments afin d'obtenir des détails relatifs à ceux-ci. Bien évidemment cet aspect peut être instrumenté au travers des mécanismes de zoom, notamment quand ceux-ci sont sémantiques. Ceci étant, l'accent est surtout mis ici sur le fait qu'une visualisation d'information, basée sur une approche graphique, par exemple, est le plus souvent moins précise qu'un texte (Lamy 2006). Aussi, dans beaucoup de situations, les composants d'une visualisation pourront nécessiter un complément d'information.

Au-delà du Visual Information-Seeking Mantra, Shneiderman introduit d'autres tâches devant être proposées à l'utilisateur dans une situation d'interaction avec une visualisation. Il s'agit du : relate, l'history, et l'extract.

Le *Relate* (l'association) doit permettre d'appréhender les relations existant entre les différents items composant une visualisation. Il s'agit ici de s'appuyer sur la vue globalisante que constitue une représentation graphique. La visualisation d'arbres ou de graphes est particulièrement adaptée à ce type de tâche. Ils ont d'ailleurs été largement exploités dans des domaines la visualisation de réseaux, qu'ils soient de télécommunication ou sociaux. Ainsi, la facilité de lecture de ce type de structure sera largement liée à la manière de les dessiner

(force direct placement, décomposition en niveaux pour les graphes orientés). Les tableaux (forme matricielle de visualisation) constituent également une forme de visualisation adaptée à ce type d'activité (Goody 1979) que l'on retrouve dans des formes très évoluées pour l'analyse de réseaux sociaux sous la forme de « Matrix Explorer » par exemple (Henry et Fekete 2006).

L'*history* (historisation) correspond au besoin de conserver une trace des actions effectuées par un utilisateur. Les systèmes interactifs permettent d'opérer de nombreuses actions simultanément sur une visualisation. Ceci induit l'affichage de nombreuses informations sous la forme d'une succession d'images. A travers l'historisation, il s'agit de pouvoir supporter le retour en arrière dans une navigation, le replay ou encore l'affichage progressif.

Enfin, l'*Extract* (extraction) correspond à la capacité de pouvoir extraire un sous-ensemble d'éléments constituant la visualisation. Au travers de cette tâche, on identifie assez bien la visualisation comme un support à l'analyse de donnée, mais également à la valorisation de ces données dans un autre contexte ou outil. Il s'agit toutefois d'implémenter ce genre de mécanismes avec précaution du fait de la notion de Gestald. Une visualisation, une représentation graphique, constitue un « tout » qui est plus que la somme de ces parties, à l'image de ce que l'on peut retrouver en sémantique textuelle (Rastier 1987). Le fait d'extraire une partie des informations constituant une visualisation les décontextualisent et peut avoir un impact non négligeable sur le sens.

### 1.3. Conception d'une visualisation

Comme nous l'avons évoqué au début de ce chapitre la visualisation d'informations doit représenter un moyen d'observer des données abstraites. Nous venons ainsi d'évoquer dans les deux sections précédentes, un panorama des techniques de visualisation, puis un aperçu des mécanismes interactifs recommandés pour accompagner la visualisation produite. Tout ceci ne constitue que des moyens, une boîte à outils, dans lesquels il est possible de choisir pour construire une interface de visualisation, mais ne représente pas en soi une méthode de conception. Une méthode de conception doit tout d'abord s'attarder sur les données abstraites à visualiser. Il s'agit en particulier de s'intéresser à leur nature.

### 1.3.a Types de données à visualiser

La conception d'une visualisation ne peut être dissociée de la nature des données que l'on souhaite représenter. Dans (Shneiderman 1996), sept types de données sont mises en évidence :

- 1-dimensional
- 2-dimensional
- 3-dimensional
- Multi-dimensional
- Temporal
- Tree
- Network

Les données de dimension 1 (« 1-dimensional ») se caractérisent par une forme d'organisation séquentielle et propositionnelle. Les données de dimension 2 (« 2-dimensional ») peuvent être illustrées par les informations spatiales et les cartes. Les données de dimension 3 (« 3-dimensional ») correspondent, par exemple, à des représentations numériques de volumes. Sans remettre en cause cette typologie, il semble qu'elle ne s'intéresse pas aux données élémentaires. En effet, les données de dimension 2 ou 3 ou encore multi-dimensionnelles (« multi-dimensional ») correspondent, en fait, au croisement de plusieurs données élémentaires qui vont constituer un tout. Le croisement constitue effectivement une nouvelle donnée, la visualisation révélant la mise en relation des différentes données. Dans le cadre de la visualisation, il s'agit donc de s'intéresser à la nature des données élémentaires pour ensuite, nous le verrons ci-après au travers de la sémiologie graphique, les associer à des variables visuelles elles aussi élémentaires.

Les données temporelles (« temporal ») peuvent être illustrées par un jeu de données traduisant une succession de dates. Ce type de donnée ne constitue en fait qu'une donnée de dimension 1, mais traduit une séquentialité dans les données caractérisée par le temps.

Les données de type arbre (« tree ») ou réseau (« réseau ») sont à l'image des données de dimension supérieures à 1. En effet, il s'agit « *d'ensemble de données pouvant être structurées de manière arborescente* » ou « *d'ensembles dont on peut extraire une structure* » (Lengler et Eppler 2007). Il s'agit d'ailleurs d'ensemble de données à 2 dimensions puisqu'un

arbre ou un graphe peut être représenté sous la forme d'une matrice d'adjacence (Lacomme et al. 2003). Ainsi, les arbres et les réseaux ne correspondent pas en eux-même à des types de données, mais plutôt à un type de représentation dont la construction permet de croiser des données élémentaires.

### 1.3.b Types de représentation

Nous avons introduit dans le début de cette section une classification des techniques de visualisation. Ces différentes techniques parfois propres à un domaine, un contexte ou encore un usage particulier, partagent certaines propriétés dans leur construction. Il est ainsi possible de lister les visualisations de type (Beaudoin-Lafon et Hascoët 2001) :

- Liste,
- Arbre,
- Graphe ou Réseau,
- Modèle vectoriel.

La visualisation de listes doit permettre d'offrir une visualisation d'un ensemble de données sur une dimension. Les techniques de visualisation sont alors nombreuses et éclectiques. Dans ce type de visualisation, la construction doit permettre non pas de mettre en valeur le croisement entre différentes données, mais bien la comparaison entre des données d'un ensemble. Il est alors possible ou d'ajouter des « *indices visuels à une représentation classique permettant la comparaison entre les éléments* » (Beaudoin-Lafon et Hascoët 2001) (cf. Figure 31) ou d'intégrer des mécanismes interactifs de visualisation. On peut, par exemple, opérer une déformation sur la visualisation pour obtenir une vue intégrant à la fois vue globale et contexte : les mécanismes de *fisheye* en sont un exemple.

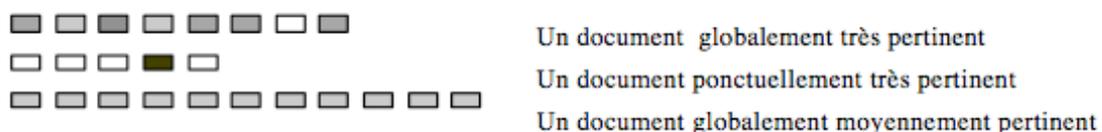


Figure 31 TileBars (Hearst 1995), visualisation de la distribution des paragraphes pertinents dans un document (source : (Beaudoin-Lafon et Hascoët 2001))

La *visualisation d'arbre* est en relation avec des jeux de données à deux dimensions. Il s'agit de visualiser des données élémentaires reliées les unes aux autres (un fils ne pouvant avoir

qu'un seul père) et ce à partir d'une racine. L'arbre est une forme de structuration très courante qui a suscité un grande fécondité dans les techniques de visualisation. Nous citerons ainsi de manière non exhaustive les listes indentées (Figure 32), les diagrammes (Figure 33), les approches surfaciques, les dessins en 3D ou encore la géométrie hyperbolique ((Figure 35).

### Racine

- Nœud 1
  - Nœud 3
  - Nœud 4
- Nœud 2
  - Nœud 5
  - Nœud 6

Figure 32 Liste indentée

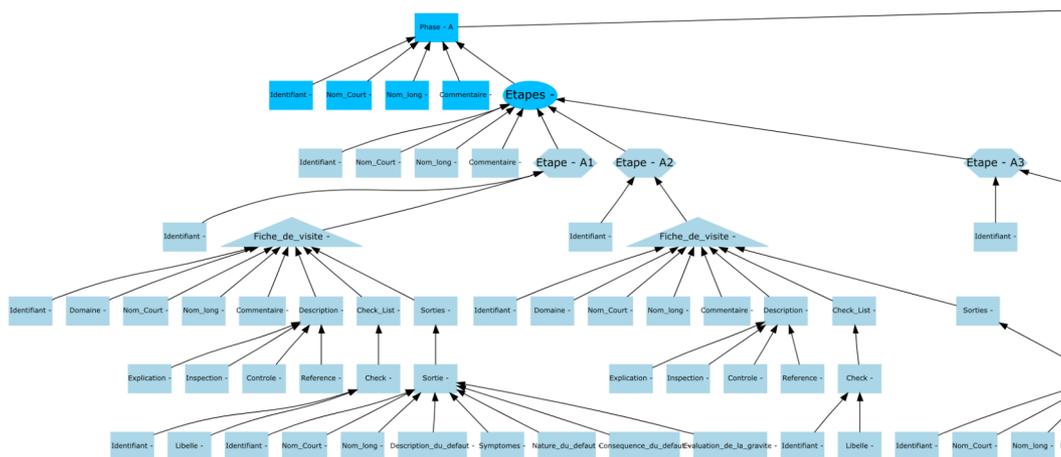


Figure 33 Génération d'un graphe orienté basé la décomposition en niveaux d'un graphe

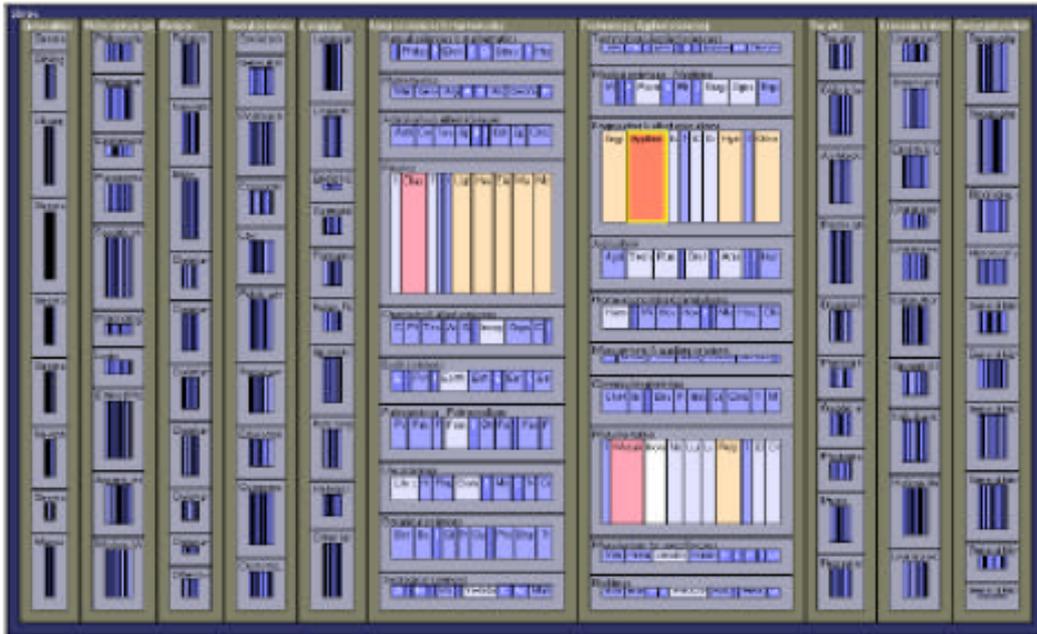
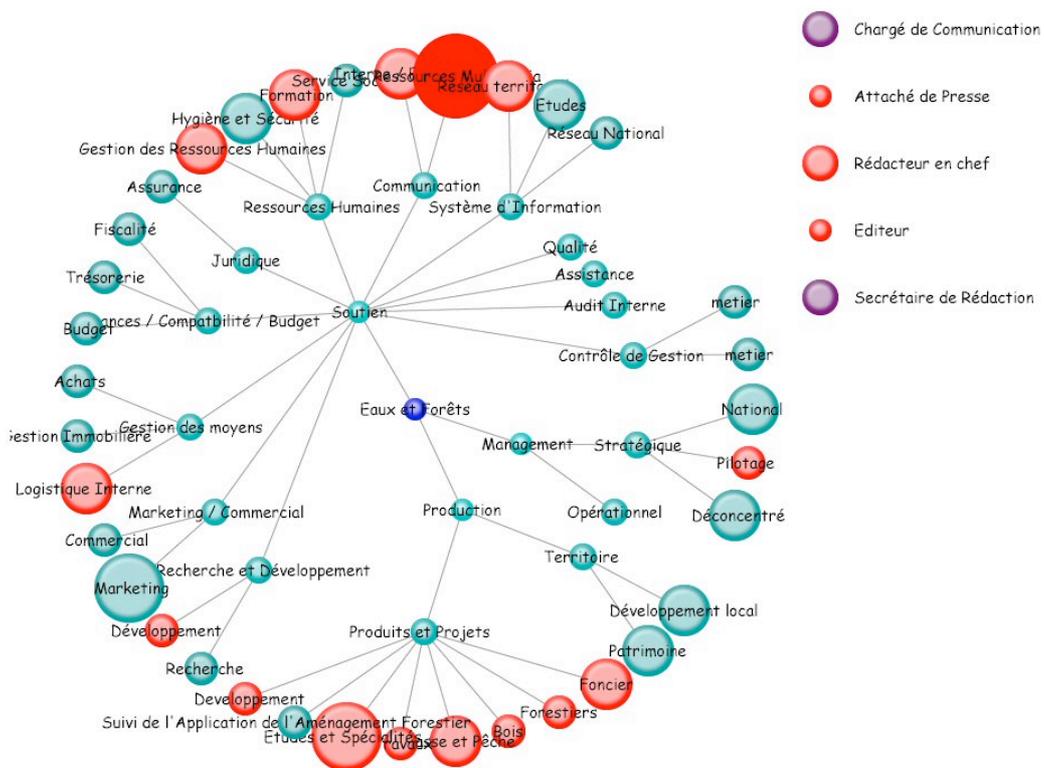


Figure 34 Approche surfacique de visualisation (ex : tree-maps (Shneiderman 1992)) d'arbre ou l'imbrication des cadres traduit la hiérarchie et la profondeur à l'image d'une poupée russe (source : (Beaudoin-Lafon et Hascoët 2001))



(Figure 35 Visualisation hyperbolique d'un arbre (source : <http://ontology.univ-savoie.fr/condillac/en/activities/research/cs/>)

La *visualisation de graphe* ou de réseau doit permettre de traduire visuellement des données relationnelles comme des modélisations entités-relations. La visualisation d'arbre que nous venons d'évoquer ne représente qu'un cas particulier de graphe. Les techniques exploitées dans ce type de visualisation renvoient aux algorithmes de construction de graphe classique (décomposition en niveau, force direct placement) ou encore le multi-arbre.

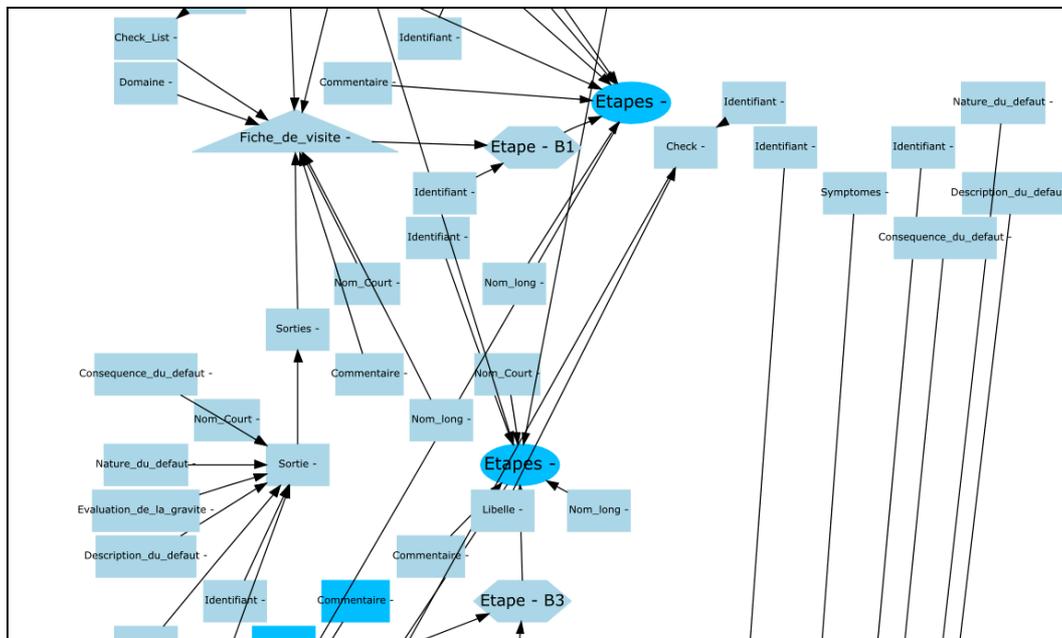


Figure 36 Génération d'un graphe orienté avec un algorithme de « force direct placement ».

La *visualisation de modèles vectoriels* renvoie à la représentation de données multidimensionnelles (nombre de dimensions supérieur à 1). Dans ce type de cas, on exploitera des métaphores de paysage, l'utilisation d'espace distance où l'on développera des cartes interactives dynamiques. Ses différentes techniques de visualisation mobilisent les dimensions du plan, éventuellement la 3<sup>ème</sup> dimension. De là, la similarité entre les données est traduite par la proximité dans l'espace.

Les différents types de construction visuelles que nous venons de présenter ne proposent encore qu'un choix de solution de visualisation qui sont plus ou moins adaptées aux différents types de données que nous avons évoqués, mais pas une méthode de construction de visualisation.

Les différentes constructions vont ainsi exploiter ou la couleur ou le niveau de couleur pour la mise en valeur (exemple des *Tilebars* proposé pour les listes) ou les dimensions du plan (arbre, graphe, modèles vectorielles) pour dissocier les différentes données et les mettre en

relation. Les différentes constructions correspondent à des systèmes de représentations bien définis sur lesquels il est possible de mapper des données. Ces systèmes de représentation mobilisent des moyens de visualisation correspondant aux variables visuelles. Ils imposent un certain contrat d'écriture et de lecture dans la représentation. La section suivante propose d'aborder le domaine de la sémiologie graphique qui, au-delà de proposer un système de représentation prédéfini, introduit une réelle méthode de conception d'image.

## 2 - La sémiologie graphique (Bertin 1973)

Bernard Morand rappelle qu'« *un beau dessin vaut mieux qu'un long discours* » (Morand 2000). Cette maxime, largement employée lorsque l'on enseigne, est évoquée dans un discours empreint de sémiotique. Le « beau » ne renvoie pas à l'esthétisme et c'est pour cela qu'il évoque le « bon », sous-entendu véhiculant le bon sens et une facilité de compréhension (Morand 2000). Le « bon » est opposé à « long ». Le dessin qui est évoqué ici renvoie à une représentation graphique, une image (Morand 2000).

La représentation graphique correspond à une image particulière. A la différence de l'image figurative, telles que celles produites dans le milieu artistique où l'interprétation ne peut être déterministe, la représentation graphique est construite pour « *retenir, comprendre et communiquer les observations* » (Bertin 1973). Bertin dans son ouvrage *Sémiologie graphique* publié en 1973 définit la graphique, comme « *un système monosémique comme les mathématiques* » mettant, en plus, à contribution la « dimension spatiale ». Il s'intéresse à l'image et à la manière de la construire.

Le caractère monosémique du langage employé doit être identifiable, soit de manière implicite, soit au travers de la définition d'une légende. L'apport d'un langage monosémique est essentiel. Il l'est en géographie, la discipline dont Bertin est issue, où la problématique de construction de diagrammes, de cartes ou encore de tableaux doit permettre une interprétation unique de l'information, ces modes de représentations devant constituer des objets permettant l'analyse, la découverte, et le traitement de l'information. Il s'agit de réduire la confusion au regard de l'image afin que les personnes souhaitant l'exploiter, les participants à une analyse, s'accordent sur le sens de ce qui est représenté, le message véhiculé par l'image. Ainsi, dans une « *situation sémiotique* » (Enjalbert 1996), une personne face à une représentation graphique, la sémiologie évoquée par Pierce, doit être déterministe mettant en relation de manière claire les signes portés sur le support de visualisation et une signification.

La graphique présentée et définie par Bertin comme un outil de traitement de l'information peut donc être définie comme « *un système de signes qui permet de transcrire les relations de différence, d'ordre ou de proportionnalité existant entre des données qualitatives ou quantitatives. Son domaine s'étend à la construction de toutes les cartes (habituellement différenciées entre topographiques et thématiques), à la construction des diagrammes (constructions matricielles et courbes) et des réseaux (organigrammes, généalogies )* ». (Bonin 1997). La graphique en tant que système de signes constitue donc un langage introduisant à la fois des signes pouvant être exploités ainsi qu'une grammaire qui régit leur association. De là, « *la graphique structure la construction des images à partir d'une grammaire qui s'appuie sur les lois de la perception visuelle, perception universelle : la sémiologie graphique* » (Bonin 1997).

La graphique doit donc être exploitée pour faire transiter un message. Au préalable de la construction de l'image en elle-même, il s'agit d'analyser l'information et les questions auxquelles on souhaite qu'elle puisse répondre en tant que moyen de découverte. L'information est ainsi vue comme le *contenu*, alors que l'image produite s'appuie sur un *contenant* se composant des moyens du système graphique.

## 2.1. Contenu et contenant d'une représentation graphique

### 2.1.a Le contenu

Le contenu correspond à l'information que l'on souhaite communiquer ou analyser. La représentation graphique en tant qu'outil de traitement de l'information doit permettre de s'interroger sur les données. L'*information* est définie en sémiologie graphique comme « *une série de correspondances entre un ensemble fini de concepts de variation ou de composantes* » (Bertin 1973). Ainsi sur un graphique présentant un axe des abscisses et un axe des ordonnées l'information se trouve dans les points définissant une correspondance entre deux valeurs des composantes portée sur ces deux axes.

Ces différentes correspondances peuvent être caractérisées par un *invariant* commun à l'ensemble des correspondances et définissant de manière complète et invariable les données représentées. Nous avons évoqué la notion de *composante* qui correspond aux concepts de variation que l'on met en relation. Les composantes vont se caractériser par leur longueur renvoyant au nombre de valeurs différentes identifiables. Cette longueur pourra être courte

(entre 2 et 10 éléments) ou longue (généralement au-dessus de quinze éléments). Les composantes courtes simplifient les problèmes de sélection et d'association.

La longueur est bien évidemment en relation avec la nature de la composante qui pourra être qualitative, ordonnée ou quantitative. Une *composante qualitative* peut être appréhendée selon des critères de différence ou ressemblance permettant la catégorisation et la classification. Par contre, elle ne peut être ordonnée de manière naturelle et implicite. Une *composante ordonnée* présente par contre un ordre universellement accepté permettant de trier les éléments. Enfin, une *composante quantitative* lorsqu'il est possible de définir des distances entre éléments, et ce, au travers d'une unité de mesure.

### 2.1.b Le contenant ou les moyens du système graphique

Dans le système graphique, les variables visuelles constituent le matériau de base, les signes, exploitables pour construire une image. Un graphique est en effet constitué de points et de tâches dans le plan, soit deux dimensions, que constitue le support d'inscription du graphique. Pour chaque point ou tâche, il existe une troisième basée sur les variables rétiniennes. Ces variables correspondent à des concepts de variation pouvant être perçus et identifiés par l'œil. On recense donc six variables visuelles auxquelles il convient d'ajouter le plan : la taille, la valeur, le grain, la couleur, l'orientation et la forme. Chaque variable va être caractérisée par sa longueur qui traduit la capacité d'une variable à représenter un nombre plus ou moins grand de divisions d'une composante.

Le *plan* correspond à l'espace à deux dimensions du médium que l'on exploite. Ainsi, il s'agira des 2 dimensions de la feuille, du tableau ou encore de l'écran. La différenciation entre les différentes valeurs appartenant à une même composante se fait au regard de leur position dans l'espace. Il s'agit de la variable la plus longue avec la forme, l'espace n'étant limité que par les dimensions du médium, du support d'inscription exploité.

La *taille* correspond à la perception de la variation de surface des tâches exploitées dans le graphique. Ainsi, il est possible de représenter des éléments du graphique selon différentes dimensions. La longueur de cette variable est relativement courte, se situant entre 5 et 20 palliers en fonction du type de perception recherché.

La *valeur* correspond à « la progression continue que l'œil perçoit dans la suite des gris » (Bertin 1973). Il s'agit du dégradé de couleur allant du blanc au noir au travers de l'utilisation des gris. La valeur est indépendante de la couleur. La Figure 37 illustre trois variations de

valeur dans trois couleurs différentes. La longueur de cette variable se traduit par six ou sept paliers, celle-ci étant tempérée par la distance entre le noir et le blanc ainsi que la taille des tâches.

Le *grain* correspond à « *la quantité de taches séparables contenues dans une surface unitaire* » (Bertin 1973). Il est possible de rapprocher cette notion de celle de texture. La variation est particulièrement adaptée et efficace dans les images imprimées. Il est par exemple possible de créer des phénomènes de vibration de la pupille sur une zone précise de l'image via une succession de rayures basée sur 50% de blanc et 50% de noir. Le grain est par contre plus difficile à introduire sur les images sur écran, les phénomènes vibratoires pouvant être amplifiés par le balayage, ce qui se révèle alors particulièrement gênant. La longueur de cette variable est largement liée à la tâche qui constitue la trame. Il sera possible d'introduire de deux à cinq paliers en fonction de l'implantation.

La *couleur* doit absolument être dissociée de la valeur que nous avons évoqué précédemment. Elle renvoie à une « *excitation colorée* » (Bertin 1973), qui provoque une différenciation pouvant être perçue entre deux plages de même valeur. Sur la Figure 37, les trois couleurs présentées sont exploitées pour illustrer la valeur. Valeur et couleur peuvent ainsi représenter un tableau de données à deux entrées.

L'*orientation* correspond à une variation angulaire d'un signe dans la représentation graphique pouvant se faire ressentir par rapport aux autres signes. Cette variation est propre à certains signes présentant des bords linéaires, puisque la rotation d'un cercle ou d'un point ne pourrait être perçue. En fonction de l'implantation, la longueur pourra aller de deux à quatre orientations différentes, voire plus en tenant compte du fait que cette variable ne permet pas de très bien différencier les différentes valeurs lorsqu'elle est trop importante.

Enfin, la *forme* constitue la variable la plus longue puisque les formes possibles ne sont limitées que par l'imagination. Elle est quasiment infinie, bien que sa longueur pourra être limitée par la présence de formes symboliques préexistantes pouvant avoir une signification dans le contexte dans lequel on souhaite diffuser une image. Ainsi, même s'il est toujours possible de définir une légende particulière, le fait d'exploiter un symbole particulier ayant une signification dans un domaine, en souhaitant changer sa signification sera source de confusion et donc pourra mener à introduire de mauvaises interprétations d'une représentation graphique.

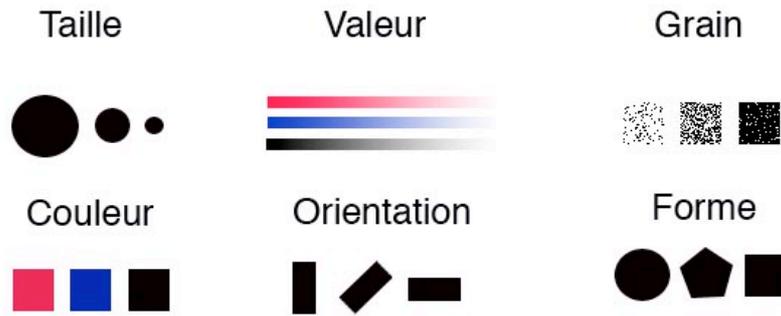


Figure 37 – Illustration des variables rétiniennes dans le système graphique

Nous proposons maintenant de nous attarder sur la méthode de construction d'images proposées par Bertin qui s'appuie sur une étude empirique de la perception humaine de ces variables visuelles.

## 2.2. Méthode de construction

Construire une image c'est « *faire correspondre des variables visuelles aux composantes de l'information* » (Bertin 1973). En ce sens, il faudra autant de variables visuelles que de composantes à faire apparaître. Ceci constitue la limite du système graphique. Le nombre de composantes caractérise également le nombre de questions qu'il est possible de se poser en analysant une représentation graphique. Toute la sémiologie graphique est basée sur une analyse empirique de la perception des variables visuelles.

Nous avons évoqué la graphique précédemment comme introduisant une grammaire sur la perception visuelle. La construction d'une image débute par un inventaire des composantes des informations que l'on souhaite faire apparaître. Il s'agit ensuite de choisir la variable visuelle la mieux adaptée pour représenter chaque composante et ceci dans un souci d'efficacité de l'image au regard du traitement de l'information.

### 2.2.a Niveaux de perception visuels et efficacité

Le fait de vouloir faire transiter un niveau de perception visuel particulier doit permettre de répondre à une question au regard de l'image.

Quatre types de perceptions sont évoqués :

- l'associative,
- la sélective,
- l'ordonnée et

- la quantitative.

La *perception associative* doit permettre de regrouper des éléments constitutifs d'une représentation graphique. Il s'agit de pouvoir répondre à la question : « est ce que tel élément (signe) est associé à tel autre ? ».

La *perception sélective* doit contribuer à isoler une catégorie d'éléments dans l'image, c'est à dire être capable du premier coup d'œil, à la première lecture de l'image, de sélectionner l'ensemble des éléments appartenant à une même catégorie, partageant la valeur d'une composante.

La *perception ordonnée* doit permettre de faire transiter une notion d'ordre permettant la comparaison et donc le tri d'éléments. Cette perception doit être directe et ne pas nécessiter de consulter une légende.

Enfin, la *perception quantitative*, associée aux valeurs numériques, doit permettre d'identifier de manière naturelle le rapport entre deux signes correspondant alors à la différence entre deux nombres. Il s'agit également de pouvoir faire ressentir les notions de distance et ce sans légende.

Bertin introduit la notion d'*efficacité* de l'image. Il s'agit, en fait, d'évaluer si une image est plus adaptée qu'une autre pour répondre à une question relative aux informations représentées. Cette notion d'efficacité se caractérise par le fait qu'il faut moins de temps avec une image qu'avec une autre pour répondre. Les travaux empiriques de Bertin ont ainsi permis de mettre en relation les variables visuelles (variables rétinienne ainsi que le plan) avec les niveaux de perception évoquées précédemment. Notons que ces travaux empiriques ont ensuite été confirmés en neurophysiologie (Lamy 2006).

Le Tableau 1 présente cette mise en relation, les « X » signifiant que telle variable est particulièrement efficace pour faire transiter tel ou tel niveau de perception visuelle. Ce tableau permet d'identifier la polyvalence de l'utilisation du plan, rendant possible la transition de l'ensemble des niveaux de perception. Il s'agit d'ailleurs de la seule variable permettant de percevoir les quantités. Nous mettrons l'accent sur les « capacités » réduites de la couleur, contrairement à la valeur. Bertin évoque d'ailleurs que la couleur est optionnelle d'autant plus qu'elle se révèle coûteuse à imprimer.

Tableau 1 – Niveau d'organisation des variables visuelles

<b>Variables</b>	<b>Niveaux de perception</b>				<b>Longueur</b>
	<i>Association</i>	<i>Sélection</i>	<i>Ordre</i>	<i>Quantité</i>	
<i>Plan</i>	X	X	X	X	Grand
<i>Taille</i>	X	X	X	X	5 à 20
<i>Valeur</i>	X	X	X		6 à 7
<i>Grain</i>	X	X	X		2 à 3
<i>Couleur</i>	X	X			10
<i>Orientation</i>	X	X			12
<i>Forme</i>	X				$\infty$

### 2.2.b Les types de représentation

Nous venons de décrire quelques règles relatives à la construction des représentations permettant de choisir une variable visuelle adaptée en fonction du niveau de perception que l'on souhaite faire transiter dans le traitement des données.

En complément, il est possible d'identifier trois types de représentation graphique qui se différencient en fonction des modes de correspondance possibles entre les composantes. Ce sont : les diagrammes, les réseaux et les cartes.

Les *diagrammes* sont caractérisés par la possibilité de mettre en relation toutes les divisions d'une composante avec l'ensemble des autres composantes. Ce type de construction tient tout aussi bien des tableaux que des graphiques présentant des courbes par exemple. Le plan est mobilisé présentant classiquement deux composantes de manière orthogonale, une troisième pouvant apparaître à la croisée des valeurs par l'utilisation de variables rétinienne. Les graphiques à trois dimensions entrent également dans cette catégorie.

Dans les réseaux, les correspondances entre composantes peuvent s'établir entre tous les éléments d'une même composante. C'est typiquement le cas de la représentation de graphes sous la forme de nœuds et d'arcs les reliant. Bertin introduit que dans les réseaux seul les informations de liaison sont significatives dans le plan et définissent des correspondances. La grosseur des points ou encore la longueur et la forme des lignes ne peuvent introduire qu'une caractérisation des entités élémentaires de la figure.

Enfin, les cartes sont présentées comme une construction basée sur des correspondances entre composantes géographiques. Le plan est alors mobilisé pour représenter l'espace géographique.

La sémiologie graphique de Bertin introduit encore de nombreux éléments à prendre en compte dans la construction des représentations graphiques, mais nous ne les approfondirons pas ici.

### 2.3. Au-delà de Bertin

La sémiologie graphique de Bertin se focalise sur la construction d'images fixes. Ceci étant Bertin évoque, dans la préface, que l'ordinateur trouvera « *son expression la plus complète et la plus puissante* ». La passage à l'image dynamique en est un exemple. Elle est en particulier rendue possible par le support numérique et la capacité de calcul de l'ordinateur qui permet d'étendre les variables visuelles. Ainsi, MacEchren propose les variables dynamiques suivantes (Blok 1997):

- *display date (or moment of display): the time at which some display change is initiated;*
- *order: the sequence of frames or scenes (a scene is a number of sequential frames with no changes);*
- *duration: the length of time between two identifiable states;*
- *frequency: the number of identifiable states per unit time;*
- *rate of change: the difference in magnitude of change per unit time for each of a sequence of frames or scenes;*
- *synchronization (or phase correspondence): the temporal correspondence of two or more time series.*

D'autres notions sont également apparues pour compléter et étendre les variables. Un exemple est la notion de *pattern* qui se positionne comme une variable visuelles de niveau supérieure. Les patterns sont alors définis comme : « *Patterns consist of elements that have shape, orientation, texture, size and arrangement* » (Blok 1997). Il s'agit en fait d'éléments graphiques qui spécifient une configuration particulière de variables visuelles . Cette notion permet par exemple de mieux prendre en compte des environnements de création d'images qui ne permettent pas d'exploiter des variables visuelles aussi élémentaires que celles évoquées par Bertin. C'est par exemple le cas lorsque l'on utilise un logiciel proposant une

bibliothèque de composants graphiques. Dans ce cas de figure, le logiciel va pouvoir introduire de nombreuses contraintes aboutissant à une marge de manœuvre limitée sur variables visuelles disponibles.

### 3 - Ingénierie de la forme, de l'interface et du document

Ce chapitre est concentré tout d'abord sur la visualisation d'information puis sur la sémiologie graphique. Nous avons pu voir que la visualisation d'information offre un ensemble d'approches ou de techniques qui proposent chacune un ensemble de moyens de visualisation et d'interaction. L'éventail de techniques, bien que caractérisées au niveau de leur construction, comme nous avons pu le voir, pâtit du fait qu'elle ne sont pas nécessairement caractérisées au niveau de leur pertinence sur le plan cognitif. De plus, leur utilisation est le plus souvent laissée à la discrétion du concepteur puisqu'elles ne sont pas nécessairement accompagnées d'une méthodologie spécifiant leur mise en oeuvre.

La sémiologie graphique, pour sa part, apporte une réelle méthode de construction d'image à partir des variables visuelles élémentaires : plan, taille, valeur, grain, couleur, orientation, forme. Il s'agit d'une méthode destinée à construire une image support au traitement de l'information, au sens où elle doit permettre de répondre à une question relative aux informations visualisées. Ainsi, une notion d'efficacité est proposée permettant de définir si une variable visuelle est adaptée ou non pour faire transiter telle ou telle information qui doit être, au préalable, rapprochée d'un niveau de perception visuelle.

La sémiologie graphique, définie sur la base d'une étude empirique, demeure une référence même si elle ne constitue pas une méthode usuelle dans la visualisation d'information. Il y a bien évidemment une certaine méconnaissance de ces travaux dans certains champs disciplinaires (mais il faut reconnaître que la sémiologie graphique peut être difficile à appliquer avec rigueur). La principale difficulté est en relation avec les variables visuelles proposées par Bertin. Il s'agit de variables élémentaires qui ne sont pas toujours employables dans un contexte réel. Par exemple, si l'on veut s'appuyer sur les infrastructures logicielles existantes, l'utilisation des variables visuelles sera nécessairement contrainte.

La sémiologie graphique nous semble toutefois une source d'inspiration tout à fait pertinente. En effet, elle propose une méthode de construction de l'image centrée sur la notion d'efficacité de l'image pour le traitement de l'information. On ne retrouve pas nécessairement cette notion dans les techniques de visualisation qui, elles, majoritairement, vont mettre à

disposition un ensemble de possibilités de visualisation et d'interaction. La sémiologie graphique offre également un réel apport sur le plan de la sémantique. La méthodologie propose de construire une image monosémique, à la manière des langages mathématiques. Contrairement à l'image figurale, par exemple le tableau du peintre, qui peut laisser la place à l'interprétation, les composants d'une image issue de la sémiologie graphique doivent avoir une signification bien déterminée. Celle-ci doit pouvoir être identifiable de manière naturelle et implicite ou bien être précisée par la mise en place d'une légende rendue explicite.

La sémiologie graphique permet ainsi de définir les modalités de construction d'une image qui constitue une interface entre un utilisateur et des informations. Cette image, construite avec l'intention d'accompagner le traitement de l'information, véhicule un message monosémique. De ce constat, on identifie assez bien le rapprochement possible entre la construction d'image, la conception d'interface informatique et notre positionnement ancré dans la sémiotique (cf. Chapitre 4 -2.2. page 92), qui est de la considérer comme une métacommunication entre concepteurs et utilisateurs (cf. Chapitre 4 -3 - page 99).

Du fait de ce positionnement, la construction du message véhiculé est essentielle puisqu'il représente un élément crucial dans le contexte dans lequel nous évoluons. Il s'agit de proposer un message qui puisse être correctement interprété par un usager. La méthodologie de construction d'images pourrait ainsi apporter un guide méthodologique dans la conception et la réalisation d'interfaces. Au delà, en tenant compte du rapprochement que nous avons introduit entre interface et document (cf. Chapitre 4 -3 - page 99), la sémiologie graphique pourrait également apporter des éléments de réponse pour la construction de documents. Il s'agira toutefois d'adapter les contributions de la sémiologie graphique à notre contexte d'étude, ainsi qu'à nos problématiques. Une des principales difficultés est, comme déjà évoqué, de tenir compte du fait que les variables visuelles élémentaires ne sont pas toujours accessibles aussi facilement que si l'on partait d'un espace vierge.

Dans la suite de notre étude, notre approche théorique sera donc centrée sur :

- Le document
- L'interface homme machine
- La sémiologie graphique

Notre positionnement est d'articuler ces trois axes en considérant qu'ils sont complémentaires et qu'ils doivent ainsi s'enrichir les uns des autres, tout en tenant compte des valeurs et des spécificités de chacun. Le *document* constitue, pour nous, à la fois notre cadre d'analyse – nous le verrons dans la partie suivante –, mais également un modèle de conception pour les interfaces informatiques. Considérant l'interface homme machine comme une méta communication, nous pensons que le document construit selon la triade forme, signe, medium représente un concept et un cadre de conception permettant de prendre en compte cette problématique de métacommunication ainsi que les choix de supports et de modalités d'inscription. La modalité d'inscription est en relation avec le support choisi et ses possibilités, mais, dans notre idée, devra être guidée par la sémiologie graphique qu'il conviendra, bien évidemment, d'adapter.

## Partie III - Analyse empirique des systèmes de gestion de connaissances

Les résultats de l'étude présentée dans ce rapport sont basés sur l'observation d'une équipe de recherche en gestion des connaissances au sein d'EDF. L'analyse empirique effectuée s'est focalisée sur les systèmes de gestion de connaissances mis en place dans le cadre de 2 projets réels. Nous définirons les systèmes de gestion de connaissances ici comme des Artefact informatiques (base de données, documents...) représentant un support à la conservation, à la diffusion et à l'exploitation dans des activités métiers d'une représentation des connaissances, au sein d'un document ou d'un système informatique.

Les systèmes de gestion de connaissances visent généralement deux objectifs complémentaires (Dieng-Kuntz et al. 2001). D'une part, la « capitalisation » des connaissances expertes détenues par un ou plusieurs spécialistes d'un domaine, et d'autre part le traitement et la restitution de ces connaissances au moyen d'une forme de représentation. Nous proposons dans ce chapitre d'étudier les systèmes de gestion de connaissances mis en œuvre au sein d'EDF avec comme cadre d'analyse le concept de document et les 3 dimensions telle que nous les avons évoqué au Chapitre 3 -2.1. page 72 :

- Medium
- Signe
- Forme

Au travers de la dimension *medium*, il s'agit de s'attacher à la dimension sociale des systèmes de gestion de connaissances en se focalisant sur les acteurs et leurs interactions. L'analyse selon la dimension *signe* se focalise sur le problème de la définition de la représentation de connaissances qui doit constituer un objet signifiant pour les personnes qui l'exploiteront. Enfin, l'analyse selon la dimension *forme* s'attarde sur les modalités de la représentation (ex : la mise en forme matérielle) des connaissances au regard de leur visualisation. Chaque partie de cette analyse est étayée par des éléments concrets issus de 2 projets réels de gestion de connaissances (cf. Tableau 2) :

- Les Guides de diagnostic : ce projet renvoie à une problématique de capitalisation des connaissances théoriques et des pratiques métiers de diagnostic liées à de grands ouvrages hydrauliques. Ce système est construit de façon à représenter un support à la conservation et à la diffusion d'une représentation des connaissances théoriques et pratiques liées au diagnostic. (cf. section Chapitre 1 -3.2.a page 43)

- Méthodes et outils d'aide à la gestion des actifs de production : ce projet a pour objet de permettre l'évaluation et l'optimisation de stratégies long-terme de maintenant et d'exploitation d'actifs de production, via le calcul d'indicateurs technico-économiques (cf. section Chapitre 1 -3.2.b page 46).

Tableau 2 Synthèse sur les outils et les ressources produites dans les 2 projets

Projet	Outil	Ressources produites
La capitalisation des pratiques de maintenance dans le domaine de l'hydraulique : le développement de guides	Microsoft Word Microsoft Excel Microsoft Visio Adobe Acrobat Professionnel	Documents : modalité graphique et textuelle
Méthodes et Outils d'Aide à la Gestion des Actifs de Production	Microsoft Word Système informatique de simulation	Base de connaissances formelle Documents de capitalisation Notes de synthèse des résultats

## Chapitre 6 - Sous la dimension medium

Un Système de Gestion de Connaissances quel qu'il soit, doit nécessairement être considéré et analysé sous une approche sociale de prime abord. En effet, un projet de gestion de connaissances nécessite la mise en place d'un dispositif particulier monopolisant un réseau d'acteurs (cf. Chapitre 2 -1 - page 49 pour une description du réseau mobilisé) qui vont devoir collaborer. Il s'agit de s'intéresser à la mise en relation de ces acteurs et en particulier les moyens mis en œuvre dans l'échange d'informations. L'analyse est illustrée ici par l'exemple des guides de diagnostic.

### 1 - Regard sur l'exemple des guides de diagnostic au travers de la dimension medium

Le développement du *guide* et des systèmes informatiques de gestion de connaissances en général fait le plus souvent référence à un contexte complexe. En effet, au-delà de la « simple » multiplicité des acteurs, l'analyse des pratiques de diagnostic sur de grands ouvrages implique nécessairement de devoir solliciter plusieurs domaines de compétence. Ainsi, dans les cas d'application du guide, il s'agit de consulter des experts en hydro-mécanique, en génie civil ou encore en contrôle commande. Nous sommes alors confrontés à un contexte multi-domaines et multi-points de vue (cf. Figure 38 et section Chapitre 2 -2.1. page 52).

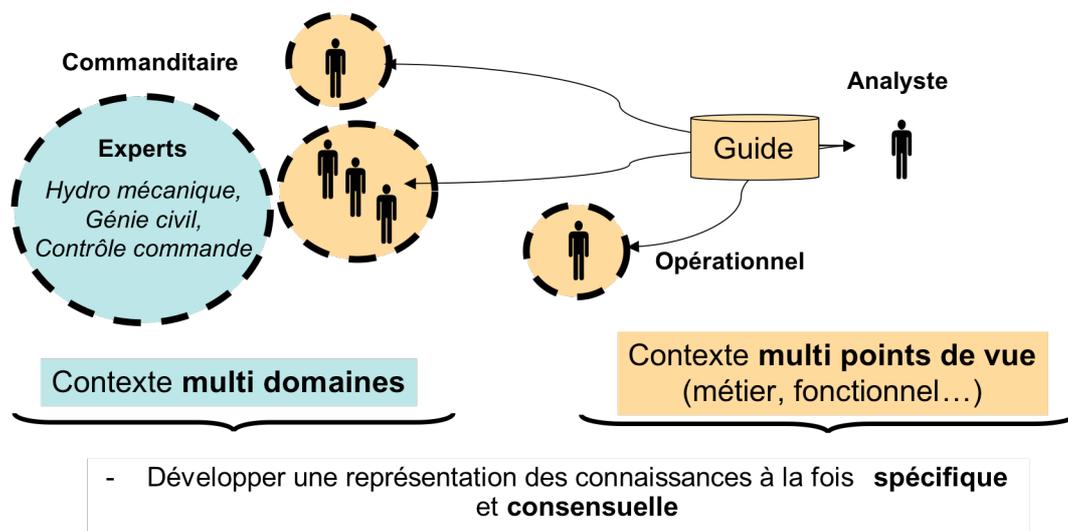


Figure 38 Contexte de développement des guides de diagnostic

### 1.1. L'analyste en médiateur

L'objectif du *guide* est d'aboutir à une description consensuelle et généraliste du diagnostic où chaque compétence sollicitée doit pouvoir se retrouver et s'identifier. Il s'agira, par exemple, dans l'activité de maintenance, pour chaque compétence, d'obtenir une bonne image de son métier au sens de bien situer ses interventions, ses actions et ses contributions. L'analyste dans sa démarche joue un rôle de médiateur entre des compétences qui amènent chacune un vocabulaire et des points de vue différents. A cela, s'ajoutent des contraintes liées aux aspects sociaux et organisationnels. Les compétences sollicitées peuvent, par exemple, introduire des clivages. Le guide, en sa valeur de médium tente de véhiculer une représentation commune qui doit alors être un point d'entrée sur la construction d'un terrain commun (Nicolle 2003). Ce terrain commun doit faciliter le dialogue et la coopération entre ces domaines d'expertise, voir même parfois les reconcilier en situation de conflit. L'analyste joue ici un rôle prépondérant dans le mise en relation des acteurs et se positionne en créateur du lien social.

Nous avons évoqué le fait que chaque domaine d'expertise introduit un point de vue différent. De même, le commanditaire peut apporter un autre point de vue et peut éventuellement se révéler intrusif dans le dispositif mis en place par l'analyste pour mener la capitalisation. De plus, il peut avoir une vision a priori du SGC à produire. Par exemple, dans une démarche de capitalisation de connaissances en rapport avec l'identification de bonnes pratiques, il pourra s'attendre à ce que l'on produise un outil prescripteur (Leplat 2004). Ceci pourra aller à l'encontre des attentes des opérationnels qui ne sont pas nécessairement demandeur d'une procédure imposée, mais plutôt d'un outil support à leurs activités, se portant plus en aide mémoire. Il s'agit donc pour l'analyste de s'adapter à la demande du commanditaire tout en trouvant parfois les arguments pour faire « évoluer » sa vision. L'analyste doit négocier entre les différents acteurs.

Les opérationnels représentent des utilisateurs qui exploiteront une restitution des connaissances dans une pratique métier qui leur est propre. Ils sont les seuls en mesure de valider le système, et particulièrement le mode de restitution, en termes d'usage et d'utilisabilité (Mallein et Tarozzi 2002; Kovacs et al. 2004). Leur point de vue est donc essentiel et doit être pris en compte dans le travail de médiation de l'analyste puisqu'il influencera nécessairement la définition du processus de restitution des connaissances.

Notons que les visions des différents acteurs sont changeantes à mesure que le projet évolue. Cette fluctuation est liée à l'interaction entre les différentes compétences, à la confrontation

des points de vue, ainsi qu'à l'impact des résultats intermédiaires du projet – pouvant faire référence aux connaissances formalisées ou encore aux moyens technologiques ou non qui sont mis en œuvre pour les présenter – lorsqu'ils sont proposés aux acteurs. L'analyste est innovateur et propose des solutions.

## 1.2. Des acteurs en interaction médiatisée

Le contexte multi-domaines et multi-points de vue inhérent à la réalisation du *guide*, situe le développement du guide comme une activité collective nécessitant une coopération forte entre les acteurs. Il est essentiel qu'un dialogue s'instaure entre eux pour converger vers un résultat consensuel. Aboutir à ce résultat peut d'ailleurs être appréhendé comme une situation coopérative de résolution de problème. Elle se caractérise par la recherche d'un équilibre négocié entre les différents acteurs en rapport avec la représentation des connaissances en terme de formalisation, mais également de restitution : le rôle de médiateur de l'analyste prend ici toute son importance.

Il doit mener une réelle négociation entre expert(s), commanditaire(s), opérationnel(s). Il est intéressant d'évoquer un instant le dispositif organisationnel de capitalisation mis en œuvre dans la problématique du *guide*. En effet, durant la phase de capitalisation, l'analyste va rencontrer dans un premier temps les acteurs de manière indépendante. A l'issue de cette première étape, il formalise les différentes visions qu'il a pu réunir durant les différents entretiens et tente d'en proposer une représentation qui les articulent. Il la proposera aux différents acteurs impliqués dans le processus, afin que chacun puisse intervenir, critiquer, annoter ce qui lui est proposé. Ce processus se poursuit jusqu'à aboutir à une stabilisation de la formalisation : une représentation consensuelle. Cette démarche illustre bien le développement du guide comme une activité collective d'un réseau d'acteurs coopérant sur la durée de manière asynchrone et délocalisée au travers d'un processus itératif et très médiatisé. Certaines réunions sont organisées afin de réunir l'ensemble des acteurs, mais elles restent rares puisqu'il peut être difficile de les rassembler dans le même temps et en un même lieu.

En outre, cette activité se traduit par de nombreuses interactions et échanges entre les acteurs, l'analyste se plaçant en pivot, en médiateur. Les échanges sont alors largement médiatisés. Les médias peuvent être observés sous plusieurs modalités puisqu'on peut observer des échanges vocaux (en « présenciel » ou non), des échanges de courriels ou plus généralement de documents qui seront autant de versions soumises à un processus « annotatif ». Dans ce contexte, le *guide*, dans sa globalité, représente lui-même un médium puisqu'à mesure de son évolution, il véhicule la formalisation et les façons de la restituer aux acteurs.

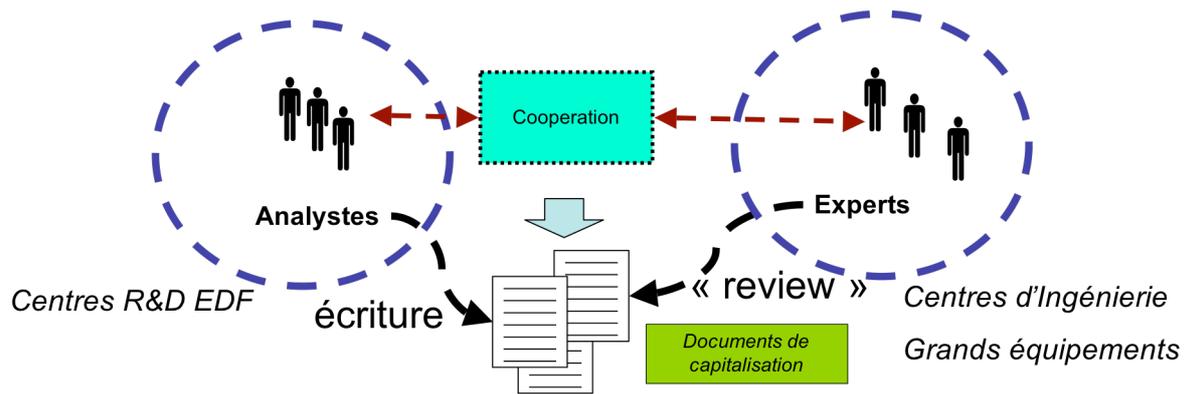


Figure 39 Contexte de capitalisation dans le cadre de la réalisation du guide de diagnostic

### 1.3. Outils exploités pour la capitalisation et la diffusion

Dans le cas du guide, le dispositif de capitalisation doit permettre d'aboutir à la rédaction d'un document qui permette de formaliser et de structurer le diagnostic relatif à l'ouvrage. L'écriture de ce document découle d'une rédaction menée par l'analyste. La production de ce document est passée par l'utilisation de plusieurs outils en commençant par le traitement texte Word, pour passer au tableur Excel et enfin se diriger vers la construction de diagrammes au travers de Visio.

Nous avons pu voir que la démarche capitalisation menée par l'analyste est centrée sur ce dernier. Dans une large mesure, c'est l'analyste qui à la fois propose et mène le processus de décision lié au choix des outils à utiliser. Ainsi, on peut observer tout d'abord que l'analyste exploite des outils relativement courants. Le choix se porte vers ces logiciels en premier lieu car ils font partie des suites bureautiques couramment déployées par les directions informatiques des entreprises. Ils constituent des outils communs sur lesquels les utilisateurs ont une certaine maîtrise : ils savent les utiliser. Le choix de l'analyste se portera alors souvent vers des logiciels qu'il sait utiliser, ou sur lesquels il peut monter en compétence rapidement. Il adaptera alors le logiciel à sa pratique, le transformant en instrument (Béguin et Rabardel 2000), pour aboutir à un logiciel auteur de représentation des connaissances.

Il faut ici mettre l'accent sur le fait que le déploiement d'une nouvelle solution peut être un exercice complexe (sans compter qu'elle a un coût financier et humain). Le fait de s'appuyer sur des logiciels disponibles facilite nécessairement la mise en place de la démarche de gestion de connaissances. En effet, ils constituent des outils communs pour les acteurs engagés auxquels ils sont habitués, pouvant ainsi favoriser l'acceptabilité de la solution de capitalisation. Son déploiement permettra également d'être l'interface de diffusion de la

représentation de connaissances sans être nécessairement confronté à des problèmes de compatibilité de format. Au travers du caractère déployé, c'est également la dimension coopérative de l'activité de gestion des connaissances que l'on appuie. En effet, les logiciels exploités étant à la fois auteur et lecteur, les experts ont par exemple l'opportunité, dans la phase de capitalisation, d'intervenir dans la représentation des connaissances (ex : utiliser le mécanisme de révision dans Microsoft Word).

## 2 - Activité de gestion des connaissances au regard de la dimension médium

Cette analyse, sous l'angle de la dimension médium de l'activité de gestion des connaissances, a permis dans l'exemple du guide de mettre en avant :

- L'analyste en médiateur
- L'omniprésence de l'échange documentaire
- Les outils exploités pour la capitalisation et la diffusion

Notons au passage que l'on peut observer des conclusions similaires dans d'autres projets. Tout d'abord, l'analyste se positionne comme un réel médiateur entre les acteurs mobilisés dans la démarche de gestion des connaissances. Cette position de médiateur prend toute son importance dans :

- La représentation des connaissances
- La définition de l'artefact à produire
- La coopération entre les acteurs

Nous avons évoqué que l'activité de capitalisation doit se réaliser dans un contexte où les acteurs sont géographiquement délocalisés. La tâche s'exécute alors pour une part de manière asynchrone et délocalisée, ce qui se matérialise par une importante médiatisation des informations identifiées et des contenus produits. Cette médiatisation se traduit en particulier par la rédaction de nombreux documents et une omniprésence de l'échange de ces documents (documents de capitalisation, cahiers des charges, compte rendu de réunion ...). Ces documents facilitent la coopération entre les acteurs du fait de leur facilité d'échange, à condition d'opter pour un format courant (PDF, Microsoft Word...). Ainsi, les documents jouent un rôle dans l'activité de médiation de l'analyste puis qu'ils constituent le médium qui

va véhiculer les informations auprès de l'ensemble des acteurs. Les acteurs sont alors dans un contexte d'interaction médiatisée via des documents qui deviennent alors des moyens de coordination (Zacklad 2004). Ils constituent des moyens de coordination, car ils véhiculent ici une représentation des connaissances sur un support et sous une forme sur lesquels les acteurs peuvent intervenir.

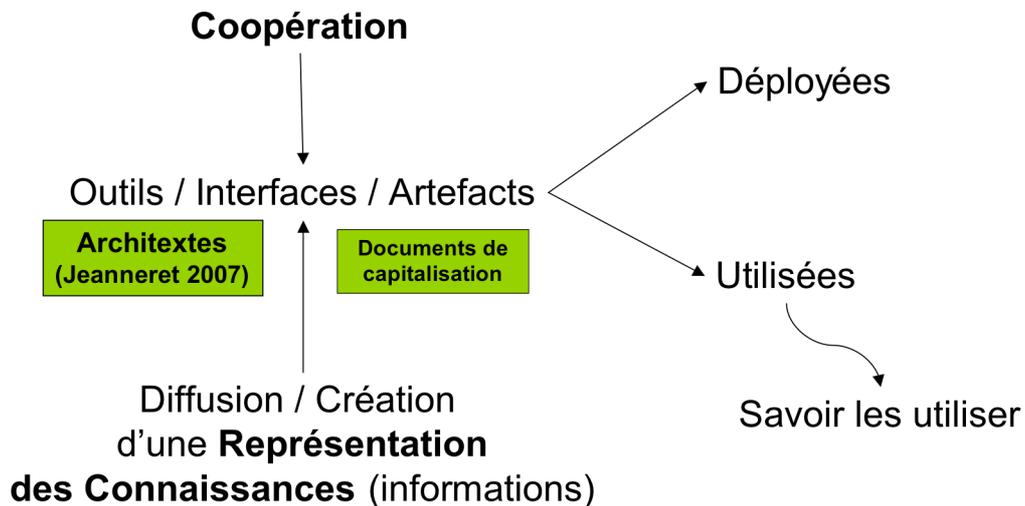


Figure 40 Les outils, interfaces et architextes exploités comme support à la coopération et la construction de la représentation des connaissances

Les documents rédigés en langage naturel représentent une interface satisfaisante entre les acteurs au prix d'une rédaction adaptée au contexte : niveau d'expertise du lecteur... Dans certaines situations, la description textuelle peut trouver ses limites en ne permettant pas de véhiculer certaines informations. Il s'agira alors de l'accompagner en introduisant d'autres type de modalités d'information tels que des images, des graphiques ou encore des vidéos. Ainsi, si le texte peut décrire la photo, quant à elle, montre. Dans le cas du développement d'artefacts informatiques, là encore le document peut ne pas suffire à décrire et expliquer tous les tenants et aboutissants. Ceci incitera alors au développement de maquettes et prototypes qui sont conçues et développées pour véhiculer, auprès des acteurs, une image de ce qu'il est possible de réaliser. De là, à l'image des documents, maquettes et prototypes constituent des médiums véhiculant une inscription sur un support auprès des acteurs. Cette vision médiums des artefact informatiques nous ramène à la vision métacommunication des systèmes que nous avons abordés dans le Chapitre 4 -2.3. page 96. Dans notre contexte d'étude, la maquette doit être développée pour appuyer une métacommunication durant le processus de gestion des connaissances entre analyste, experts et utilisateurs.

Cette métacommunication entre les différents acteurs se fonde sur l'échange d'un message. Nous avons introduit dans l'approche théorique, l'importance de bien clarifier le niveau sémiotique que ce soit dans les documents (rendre explicite le contrat de lecture et d'écriture), au niveau des IHM ou encore dans la diffusion d'une visualisation d'information. Notre analyse empirique porte alors un regard sur les systèmes de gestion de connaissances selon la dimension *signe*. Il s'agit de s'intéresser à la construction de l'objet signifiant devant transiter entre les acteurs impliqués dans la démarche.

## Chapitre 7 - Sous la dimension signe

Nous venons de porter un regard sur l'activité de gestion des connaissances selon la dimension medium. Ainsi, documents, systèmes informatiques produits et outils exploités dans ce type de démarche peuvent être appréhendés comme des médiums de la coopération entre les acteurs de par la représentation des connaissances qu'ils véhiculent. Cette représentation doit alors constituer un *objet signifiant* interprétable par ces acteurs. La construction de cet *objet signifiant* résulte en partie du processus de formalisation des connaissances où l'analyste joue un rôle prépondérant. En effet, il apporte un savoir faire de modélisation et conduit largement à la définition de l'*objet signifiant* en question. Ce savoir peut se faire ressentir dans la rédaction d'un document où il pourra apporter un esprit de synthèse, ou dans les fonctionnalités identifiées dans le cadre de l'utilisation d'un système à base de connaissances informatiques. Notons bien que la formalisation relève toutefois d'une co-construction entre les acteurs. L'analyste ne peut être lui-même un expert de l'ensemble des domaines auxquels il est confronté. Ceci nous amène à introduire qu'au-delà de son rôle de médiateur, il est accompagnateur dans la formalisation.

Ainsi, plutôt que d'analyser les acteurs et les informations qu'ils fournissent, il se situe plutôt dans une activité de « mise en analyse ». Les acteurs impliqués dans la capitalisation, et notamment les experts, sont alors incités à s'interroger sur leur propre savoir et savoir-faire. Le rôle de l'analyste ne se limite toutefois pas à la mise en analyse. En effet, le savoir-faire de modélisation de l'analyste prend parfois une grande importance face à un outil informatique imposant une saisie complexe des informations. Au-delà de la modélisation conceptuelle, il s'agit pour lui de faire valoir ses compétences d'utilisation d'un outil nécessitant une expertise particulière. Cette expertise peut être liée à son utilisation, que ce soit au niveau des fonctionnalités qu'il propose, ou encore au regard des fondements théoriques qui le supportent (ex : techniques statistiques ou probabilistes). Il prend également une place importante dans la valorisation des résultats qu'il produit. Nous proposons d'aborder la construction de l'objet signifiant devant être diffusé aux acteurs dans les 2 exemples mentionnés en début de chapitre.

## 1 - Définition du signe dans le développement des guides de diagnostic

Dans l'exemple des guides de diagnostic, la formalisation prend la forme d'une structuration, d'une indexation, d'un contenu informationnel sous la forme d'un document (sans aucune traduction dans un langage informatique formel). Ce travail de formalisation est, pour une part, fait de manière asynchrone, c'est-à-dire que l'analyste structure le guide en dehors des entretiens pour ensuite confronter les experts au document produit. Le travail est itératif, cyclique, introduisant une tâche d'annotation par les experts dans les versions intermédiaires produites.

### 1.1. Formalisation multi-domaines et multi-points de vue

Les aspects multi-domaines et multi-points de vue sont un enjeu de premier plan dans la définition de l'objet signifiant. En effet, il incombe à l'analyste de comprendre, dans une certaine mesure, les informations recueillies au sein des différents domaines de compétence consultés. Ceci va lui demander de se constituer une certaine expertise sur ces domaines pour être en mesure de communiquer avec les experts. La formalisation réalisée en collaboration avec les acteurs doit produire une représentation pivot où chacun peut, d'une part, s'identifier dans le contenu et, d'autre part, identifier les contributions des autres compétences, tout en se positionnant par rapport à elles. Il s'agit là d'un exercice difficile qui souligne encore une fois le rôle de médiateur de l'analyste, mais également le rôle de médium du document produit en tant que support véhiculant de l'information, dans la coopération entre les acteurs.

Dans la réalisation du guide de diagnostic, nous sommes confrontés à ce contexte d'étude. Il s'agit en effet de réaliser une base de connaissances (un document) offrant un panorama général du diagnostic sur l'ouvrage auquel on s'intéresse et qui pourra, a posteriori, être particularisé en fonction des spécificités de chaque ouvrage. Le *guide* doit donc refléter les informations de chaque domaine, les structurer en identifiant les interfaces entre chacun d'eux pour souligner leur complémentarité et soutenir leur coopération. La complexité inhérente à l'aspect multi-domaines est augmentée par le caractère multi-points de vue de la problématique. Les différents points de vue sont en relation avec l'aspect multi-domaines – il peut exister des divergences inter-domaine, mais également intra-domaine – mais pas uniquement. En effet, dans l'exemple du guide, il s'agit de traiter et de mettre en relation 2 modes de représentation des connaissances. Le premier est lié à une *formalisation de la pratique métier* du diagnostic qui renvoie à l'usage quotidien des experts et à la façon dont ils

appréhendent leur activité. Il en résulte une représentation séquentielle, parfois descriptive, d'étapes imbriquées dans des phases de diagnostic. Un exemple de contenu est présenté sur la Figure 41.

<b>Mode Opérateur</b>
<p><b>Inspection</b></p> <p>État de la charpente constituant le corps de bouchure :</p> <ul style="list-style-type: none"> <li>- Inspection des parties métalliques extérieures : oxydation, peinture;</li> <li>- Inspection des caissons d'allègements de manoeuvres des portes;</li> <li>- Inspection des boulons et des rivets, des déformations éventuelles.</li> <li>- Inspection des poutrelles, croisillons, entretoises;</li> <li>- Inspection de l'état de surface: peinture, corrosion;</li> </ul> <p>État des caissons :</p> <ul style="list-style-type: none"> <li>- État de l'étanchéité des caissons inférieurs;</li> <li>- État de fissuration et de corrosion des caissons;</li> <li>- Contrôle des épaisseurs;</li> <li>- Contrôle du dispositif d'accrochage;</li> <li>- Contrôle des défenses en bois et des articulations.</li> <li>- Contrôle de l'aspect extérieur du tablier : peinture, rivets, oxydation, chocs;</li> <li>- Noter les déformations éventuelles.</li> </ul> <p><b>Contrôle</b></p> <p>Contrôle des jeux de fonctionnement :</p> <ul style="list-style-type: none"> <li>- Contrôler les patins d'appui sous pression</li> <li>- Vérifier les états de surface et tenue de garniture et d'usure</li> <li>- Contrôler les freins et les garnitures</li> </ul>

Figure 41 Le point de vue métier

Le second type correspond à la définition d'une Analyse des Modes de Défaillance de leurs Effets et de leur Criticité (AMDEC). Cette AMDEC est *une représentation construite selon un découpage fonctionnel* de l'ouvrage qui est liée à sa décomposition matérielle Figure 42. La constitution de l'AMDEC représente un effort intellectuel important pour les experts puisque différente de leur pratique usuelle. Ce changement de point de vue, introduit par l'AMDEC, est alors tout à fait intéressant, car il oblige les experts à beaucoup s'interroger sur leur savoir.

## 1.2. Le document comme support de capitalisation

Dans cette formalisation multi-domaines et multi-points de vue, le document va être un carrefour et un point d'ancrage pour la modélisation. Il constitue tout d'abord le support d'inscription, le médium tel que nous avons pu l'évoquer dans la section précédente. Ce support d'inscription sera partagé par les différents acteurs pouvant intervenir dans sa rédaction. Dans le cas présent, l'analyste reste l'auteur principal puisqu'il est en charge d'articuler à la fois la représentation des différents domaines de compétence et des différents points de vue.

Dans la situation présente, il s'agit de rédiger un document de capitalisation qui se caractérise par une structuration très forte des informations qu'il contient. Cet aspect est également influencé par le caractère technique des informations transcrites. Nous nous situons alors dans un niveau de langue et un registre d'écriture particulier, où il s'agit d'identifier les informations élémentaires d'une manière très précise. Sur l'exemple présenté sur la Figure 41, des titres tels que « Inspection » et « Contrôle » véhiculent une sémantique très importante concernant les contenus préfixés. Ainsi, par exemple, l'inspection renvoie à une observation visuelle. Le contrôle pour sa part, même s'il est visuel, va introduire une notion de mesure et donc le besoin d'un instrument particulier et l'existence d'un système de référence. Notons bien ici que dans cette activité d'identification et de structuration d'entités élémentaires, l'analyste se place dans une tâche d'indexation d'informations où il délimite et identifie des zones de texte dans un corpus construit en collaboration avec les experts. L'indexation se fait en même temps que la création du corpus et se révèle très précise puisque opérée de manière manuelle. Cette indexation définit à la fois des concepts tels que l'inspection et le contrôle ainsi que la mise en valeur dans le document par la mise en forme matérielle : une mise en gras. La typographie est ainsi exploitée pour délimiter ces entités élémentaires qui constituent le diagnostic.

<p><b>Sous-fonction</b></p> <p><b>Permettre le passage des réseaux (C1)</b></p> <p>Modes de Défaillance - déformation de structure</p>
<p>Causes - perte des caractéristiques matériau</p> <p>Composant : structure</p>
<p><b>Conséquences - perte réseau (gaz à Vogelgrün ? À vérifier) ( N° fiche : 94)</b></p> <p>Indicateurs : Sécurité : 02 - Coût : 03 - Réglementation : 03 - Gravité : 3.0 - Fréquence : 01 - Criticité : 03</p> <p><b>Détection :</b></p> <p><a href="#">S_E.12.A-GC : Evaluation du risque de défaillance - Structures amont du sas - page 40</a></p> <p><a href="#">S_E.13.A-GC : Evaluation du risque de défaillance - Structure aval du sas - page 51</a></p> <p><a href="#">S_I.7.A-GC : Evaluation du risque de défaillance - Bajoyer - Sas chômé - page 139</a></p> <p><a href="#">S_I.10.A-GC : Evaluation du risque de défaillance - Structures amont du sas - Sas chômé - page 121</a></p> <p><a href="#">S_I.11.A-GC : Evaluation du risque de défaillance - Structure aval du sas - Sas chômé - page 131</a></p>

Figure 42 L'AMDEC

La rédaction du guide de diagnostic est également caractérisée par une mise en relation très forte des différentes briques élémentaires. Cette mise en relation passe à la fois par des renvois dans le texte sous la forme de références et co-références. Elles peuvent se matérialiser en langage naturel sous la forme d'anaphores, mais également par des mécanismes de lien hypertexte. Cette mise en relation dans le guide se fait particulièrement ressentir à l'articulation des points de vue. Elle représente un enjeu crucial puisque l'AMDEC véhicule une cotation du risque qui va être intégrée à la vision métier. Sur la Figure 42, les

liens hypertextes soulignés en bleu ont été introduits dans l'AMDEC et font référence à des entités élémentaires de la vision métier. Ce mode d'écriture hypertextuelle doit permettre d'aboutir à une mise en relation de la vision métier et de la vision fonctionnelle à l'image de la Figure 43.

Cette articulation des points de vue qui se révèle assez aisée lorsque le contenu du document est petit, pose très vite des difficultés lorsqu'il s'étend. En effet, la vision métier se découpe en une centaine de fiches de visite et l'AMDEC en 45 modes de défaillance. La combinatoire est alors vaste et il peut être difficile de maintenir l'articulation des points de vue cohérente manuellement au fil de l'évolution du document.

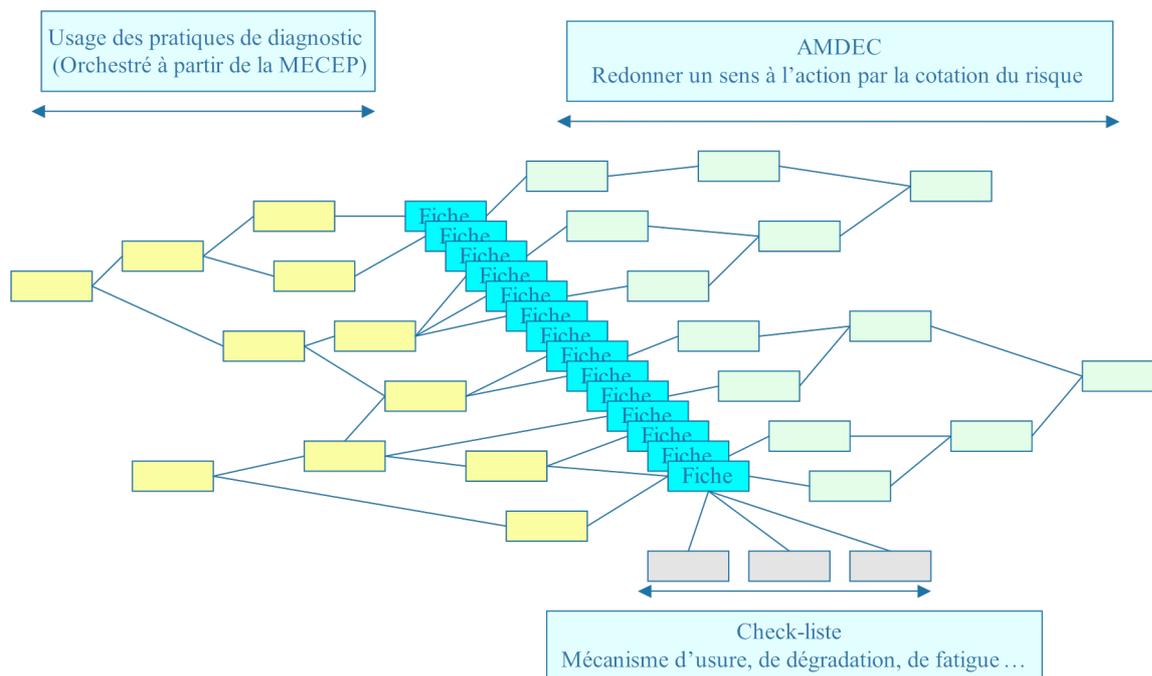


Figure 43 Le guide de diagnostic comme une articulation de points de vue

Le document, dans sa rédaction, constitue, pour l'analyste, un réel espace de conception à la fois du langage de représentation et de la capitalisation, puisque le langage se construit au travers de la capitalisation elle-même. Il est à l'image du diagramme en conception logiciel un réel espace de modélisation (Morand 2004). Ce document, au-delà de représenter un espace de conception pour l'analyste, correspond également à un réel espace de négociation et de médiation entre les acteurs de l'objet signifiant en cours de réalisation. Le document constitue un médium véhiculant un objet signifiant qui tend à coordonner les acteurs (Parfouru et al. 2006). Durant son développement, et particulièrement la phase d'hyper-rédaction dans laquelle sont impliqués l'analyste et les experts, de part leurs annotations, le document constitue un *DopA* jusqu'à aboutir à une certaine stabilisation.

## 2 - Analyse de méthodes et outils d'aide à la gestion des actifs de production sous l'angle du signe

Le travail en intelligence artificielle et le développement des systèmes experts (Hatchuel et Weil 1992) ont bien montré la difficulté qu'il y a à représenter des informations dans un langage formel. Le langage formel est ici appréhendé comme un langage codifié pouvant être interprété par la machine et ainsi participer à une logique de calcul : la formalisation est computationnelle.

### 2.1. Formalisation computationnelle

La traduction d'information dans un langage formel va devoir s'appuyer sur un utilisateur ayant une connaissance et une expertise en rapport avec ce langage. Classiquement, il s'agira de l'*analyste* qui va tenter dans le cadre d'entretien avec les experts de capter une partie de l'expertise. La capitalisation introduit ici une démarche d'entretien d'expert (Figure 44 - repère 1) , ainsi qu'une phase de traduction des informations identifiées sous une forme particulière dans un langage particulier (Figure 44 - repère 2). La traduction des informations à dire d'expert vers le langage formel constitue un enjeu majeur puisque, de la qualité de cette traduction, va dépendre la qualité des résultats produits. Il s'agit encore aujourd'hui d'un biais important dans nombre de systèmes puisque les experts ne sont pas toujours en mesure de valider la forme de représentation formelle, n'ayant pas l'expertise nécessaire en relation avec le système, son fonctionnement et les techniques statistiques et probabilistes, dans le cas présent, sur lesquelles il se fonde.

Ainsi, sur la base de ces entretiens qui font le plus souvent l'objet d'une synthèse sous la forme de documents textuels écrits, et qui ont pu faire l'objet d'un processus de relecture et de validation par les experts, l'analyste va formaliser les informations recueillies en adéquation avec les contraintes du système. Ces contraintes sont liées au langage formel, à l'algorithmique, etc... Ce processus de formalisation le met dans une situation d'interprétation (une situation sémiotique (Enjalbert 1996) où il va devoir traduire une représentation majoritairement textuelle dans le langage formel, ayant pour but de constituer une ressource exploitable par un algorithme. On voit bien ici le rôle pivot et crucial que prend l'analyste dans le processus de représentation de l'information : son expertise prend une place importante et le passage du texte au formel n'est dans les faits que peu tracé.

Il s'agit bien entendu d'une tâche difficile que de capter l'ensemble du processus de formalisation. Ceci étant, il serait certainement possible, dans un premier temps, de relier tel

fragment textuel à tel élément de modélisation en traçant par exemple les intentions (Nanard et Nanard 2005; Tazi et al. 2006) au travers d'annotations (Lortal et al. 2007; Zacklad 2007).

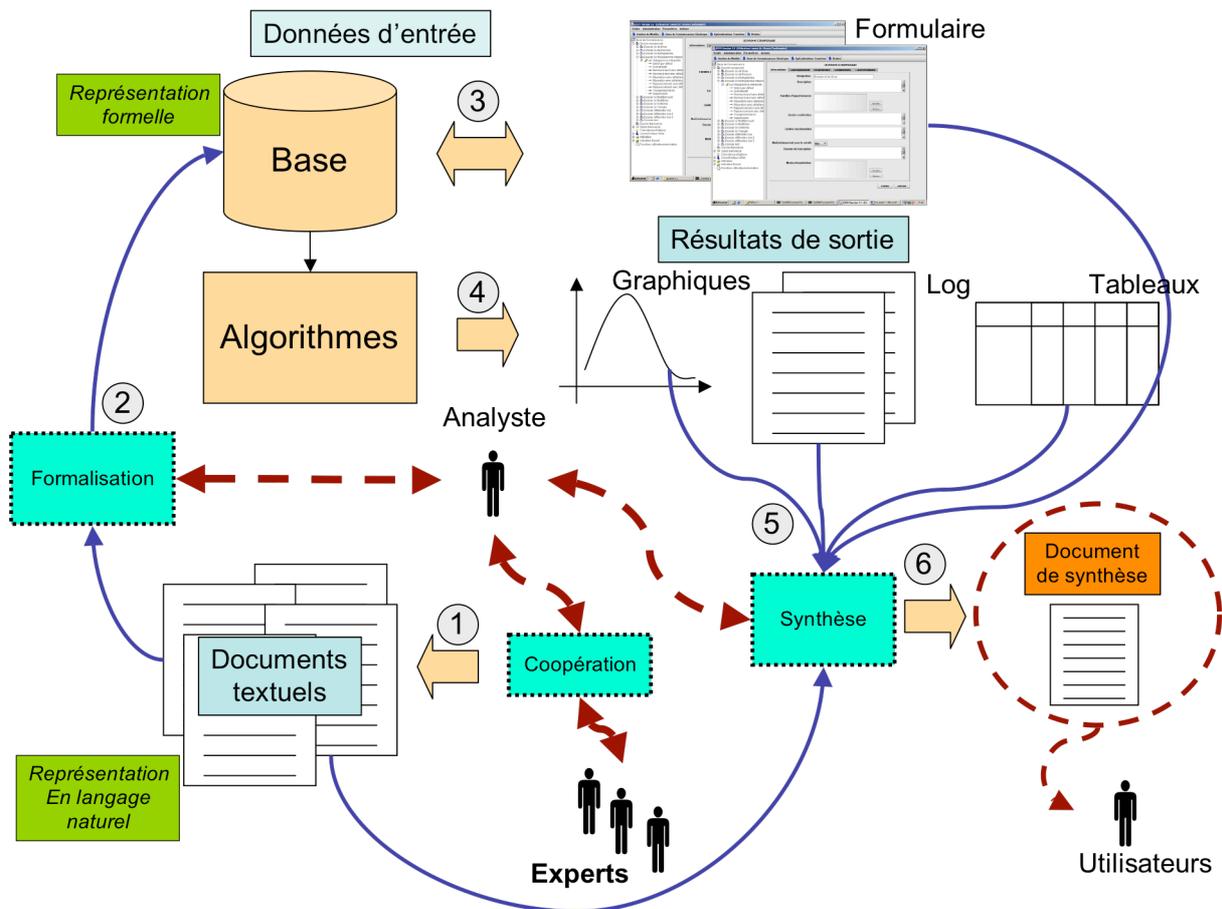


Figure 44 Contexte Socio technique dans le projet relatif à l'aide à la gestion d'actifs de production

## 2.2. La mise à disposition des résultats de sortie du système

L'analyste prend également un rôle pivot dans la mise à disposition des résultats produits à l'issue de l'exploitation de la représentation formelle par un algorithme. En effet, la lecture des résultats nécessite souvent une certaine expertise relative au fonctionnement même du système formel exploité – et par exemple les méthodes de calcul implémentées – pour les produire. Il est tout à fait admis que le système informatique ne donne que le résultat d'un calcul. Ce calcul doit nécessairement faire l'objet d'une interprétation, d'une explication (Baker 1996), correspondant à une analyse qualitative des résultats. Cette analyse qualitative se matérialise généralement par la rédaction d'un document de synthèse (Figure 44 – repère 6) qui tendra à mettre en relation les résultats de sortie (Figure 44 - repère 5) d'un système avec les données d'entrée formelles et textuelles (Figure 44 - repère 1). Ce document constitue

alors une articulation entre de nombreuses sources de données dissociées. Ces différentes sources correspondent à la fois aux documents de capitalisation produits en collaboration avec les experts, à la modélisation formelle accessible via des formulaires, aux résultats de sortie disponibles sous la forme de courbes et de tableaux, et enfin, aux traces d'algorithmes. L'exercice de rédaction met donc en relation ces différentes informations disponibles sous forme numérique en y ajoutant une argumentation qui est, pour une grande part, du ressort de l'analyste.

La rédaction de ce document de synthèse semble essentielle pour produire une réelle aide à la décision et met en avant la place de l'homme au niveau de ce genre de système. Ceci rappelle le besoin, dans le développement de systèmes informatiques, de définir en même temps les hommes (rôles, compétences, ...) qui devront accompagner son déploiement (Boullier 2001). Ainsi, ces hommes doivent posséder ou acquérir une certaine expertise. Plus globalement, le développement de systèmes formels pose toujours des difficultés tant au niveau de l'écriture cadrée dans un langage formel que dans la lecture des résultats produit (Figure 44 - repère 3 et 4). Tout ceci souligne l'importance des documents qui gravitent autour de ce genre de système. Leur exploitation nécessite un fort travail de documentation de la part des concepteurs et des utilisateurs qui possèdent une expertise sur ce système. Classiquement, cette production de documents n'est que rarement intégrée au système formel et se fait donc en dehors : cela introduit une dissociation entre le formel et le texte. Ainsi, les systèmes interactifs d'aide à la décision tendent plutôt à correspondre à des systèmes support à la construction d'une aide à la décision. La construction de cette aide à la décision correspond à la mise en relation de sources de données dissociées qui vont être collationnées sous la forme d'un document, ce dernier constituant un bon support de communication entre les acteurs impliqués dans la démarche de gestion des connaissances.

### 3 - Activité de gestion des connaissances au regard de la dimension signe

Au travers de l'analyse sous l'angle de la dimension signe, le document rédigé reste omniprésent, que ce soit dans l'exemple des guides de diagnostic ou dans le contexte de mise en œuvre d'un système à base de connaissances.

### 3.1. Face à la réalisation d'un document

Dans la réalisation des guides de diagnostic, le document de capitalisation produit doit articuler une représentation des connaissances relatives à un ouvrage hydraulique. Il se caractérise par :

- Un contenu très fortement structuré
- La définition d'entités élémentaires qui délimitent le contenu
- La mise en relation des entités élémentaires

L'analyste joue un rôle prépondérant dans la réalisation de ce document, puisqu'il construit le contenu avec les experts, tout en le structurant au fur et à mesure en définissant des concepts qui vont délimiter les entités élémentaires ; l'analyste opère ainsi une indexation manuelle du contenu. Le document peut être appréhendé comme un espace de modélisation dont le contenu évolue au fil de la collaboration entre expert et analyste : durant cette phase d'instabilité, il devient un Document pour l'Action (cf. Chapitre 3 -3 - page 77). Lorsque le contenu se stabilise, l'organisation de ces entités élémentaires, ainsi que leur mise en relation, permet d'aboutir à un modèle de structuration du document, particulièrement adapté à la représentation de diagnostic de maintenance. Ce modèle qui correspond, en fait, à la trame de rédaction du document, peut définir à la fois un contrat de lecture et d'écriture et pourrait alors être ré-exploité dans la réalisation de guide en rapport avec de nouveau type d'ouvrage. Bien évidemment, un modèle correspondra à une certaine granularité de capitalisation ainsi qu'à un certain point de vue.

### 3.2. Face à un système à base de connaissances

Dans un contexte de système à base de connaissances, il semble que le document rédigé reste encore l'interface privilégiée dans l'activité de capitalisation et de restitution. En effet, les systèmes à base de connaissances peuvent introduire une complexité que ce soit au niveau de la saisie des informations, mais également dans l'analyse des résultats produits. Le dispositif de gestion des connaissances mis en place s'appuie alors souvent sur des documents comme intermédiaires entre le système et les acteurs n'ayant pas l'expertise nécessaire à l'utilisation du système.

Au niveau de la capitalisation, les contraintes de formalisation liées aux structures de données du système ne sont pas toujours aisées à comprendre et à assimiler. Il est alors difficile de vouloir confronter tout un chacun à un système à base de connaissances. La situation la plus fréquente est alors de rédiger un document de capitalisation avec les experts pour ensuite

formaliser les informations contenues dans ce document au niveau du système. C'est classiquement l'analyste qui effectue le passage du document à la saisie cadrée par le système. Ceci pose alors le problème de l'ajustement, de l'évaluation, de la représentation des connaissances qui se font souvent entre analyste et expert sur le document et non sur la représentation formelle dans le système. Notons au passage que les documents de capitalisation produits sont, comme évoqués précédemment, très structurés, en particulier par la mise en forme matérielle (utilisation de tableau, mise en valeur du texte) et la structure logique du document (découpage en parties, sections, ...).

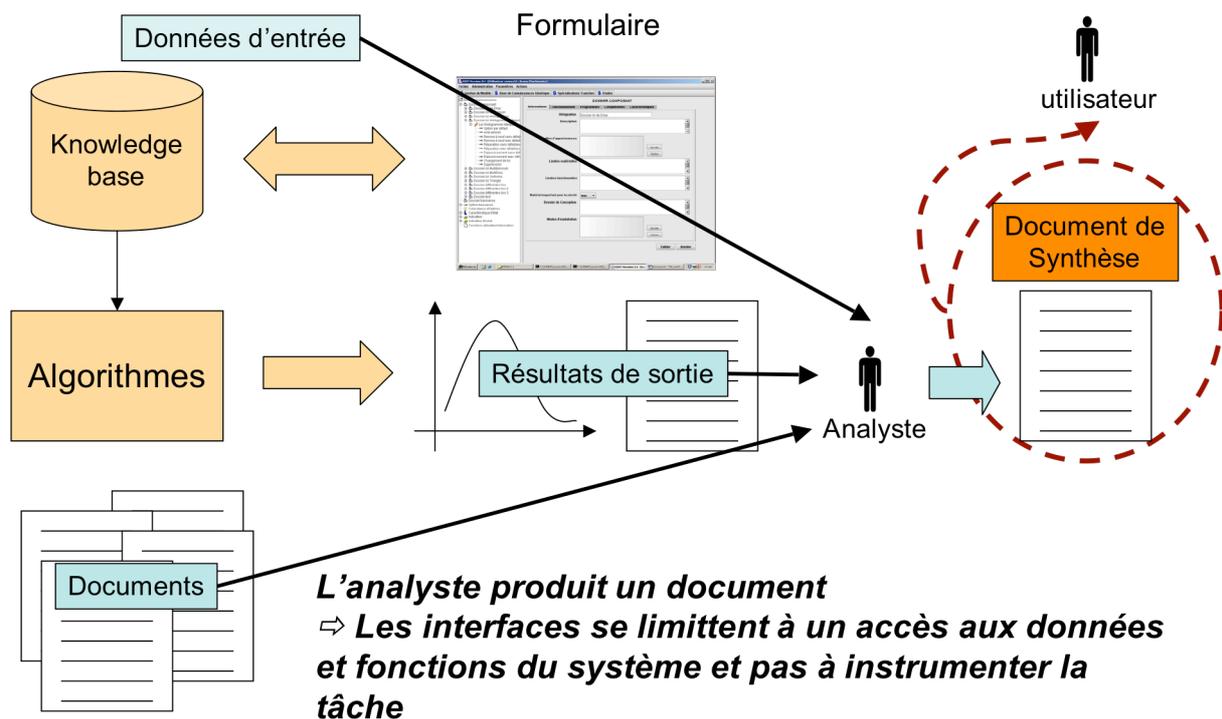


Figure 45 Restitution documentaire dans le cadre d'un système à base de connaissances

Au niveau de la restitution des résultats, là encore, l'expertise relative au système et à son fonctionnement est essentielle. En effet, l'interprétation des résultats produit doit nécessairement s'appuyer sur une compréhension du fonctionnement des algorithmes implémentés et des fondements théoriques sous-jacents. Par exemple, face à un système mettant en œuvre des algorithmes basés sur la simulation de Monte Carlo, il est difficile de demander comme pré-requis aux utilisateurs de connaître les théories probabilistes et statistiques sous-jacentes. De plus, les interfaces mises en œuvre dans ce type de système se contentent souvent de permettre une visualisation des données d'entrée et des résultats de sortie de manière dissociée. Il est alors difficile de construire une vision de synthèse qui est de toute façon difficile à imaginer a priori, en dehors d'une connaissance précise de l'utilisateur

caractérisé par son expertise, son domaine de compétence, etc... Dans ce type de situation, il est alors courant d'observer une restitution aux utilisateurs finaux sous la forme d'un document de synthèse permettant de produire une vision intégrée et interprétative des nombreuses sources d'informations, rédigée par l'analyste (cf. Figure 45). Ce document va mettre en relation, sous une forme argumentée, les résultats de sortie du système et les données d'entrée qui sont à leur origine. C'est bien l'expertise de l'analyste qui lui permet de produire un tel document qui va expliquer les résultats obtenus en intégrant implicitement la logique inhérente au fonctionnement de l'algorithme exploité.

On voit ici toute l'importance que prennent la rédaction des documents dans la capitalisation et la restitution. Elles constituent à la fois une interface entre les acteurs engagés dans la démarche de gestion de connaissances, mais également une interface entre des acteurs et un système à base de connaissances. Le document représente donc un support ou un médium, qui va véhiculer un message, l'objet signifiant, entre les acteurs. Il s'agit maintenant de s'intéresser aux modalités d'inscription de ce message sur le support. En effet, le contenu des documents s'appuie à la fois sur les mécanismes de la langue, mais également sur la mise en forme matérielle.

## Chapitre 8 - Sous la dimension forme

La définition des modalités visuelles de restitution constituent une opération lourde de sens et de signification (Eco 1988; Pédaque 2006) du point de vue de la responsabilité « autoriale » ou plutôt « éditoriale ». En effet, les aspects multi-domaines et multi-points de vue introduisent différents systèmes de signes (Eco 1988) qu'il faut tenter de faire cohabiter dans la restitution afin que les acteurs puissent l'interpréter correctement. Le SGC doit donc être construit comme un *objet de communication* : il doit appuyer la meta communication entre concepteurs et utilisateurs, entre experts et opérationnels via l'analyste.

### 1 - Analyse des documents produits

Comme nous avons pu le voir, les projets de gestion de connaissances s'accompagnent de la rédaction de nombreux documents représentant souvent des éléments clés de la démarche. La production de ces documents s'appuie, comme nous l'avons justifié précédemment, dans l'analyse selon la dimension medium, sur les environnements informatiques pré-installés par les directions informatiques. Il s'agit alors d'une utilisation de l'informatique bureautique, en particulier la suite Microsoft Office : Word, Excel, Powerpoint. Nous proposons d'analyser successivement des exemples de documents produits via Word, puis Excel, en essayant de dégager les pratiques liées à l'utilisation de ces outils.

#### 1.1. Word et la production d'un contenu orienté présentation

Le traitement de texte Microsoft Word reste aujourd'hui un standard en entreprise même si, l'on peut observer que de plus en plus d'organisations et de gouvernements tentent l'aventure du monde libre par le déploiement d'Open Office (<http://fr.openoffice.org/>). Il offre un espace rédactionnel tout à fait classique sous un mode page et permet d'associer des modalités textuelle et graphique. L'espace de rédaction est infini. Il est possible d'y intégrer des moyens de navigation au travers de l'hyper-textualité, ce qui permet de passer outre un simple mode d'accès linéaire à l'information.

L'observation des contenus nous permet d'identifier une structuration importante des informations. Cela est en particulier dû à l'activité de capitalisation des connaissances que nous rapprochons de l'activité d'indexation (cf. à l'analyse selon la dimension signe). Le

contexte technique dans lequel nous évoluons influence également la rédaction. Cette structuration s'appuie sur l'utilisation de la mise en forme matérielle. Le découpage logique du document en parties, sections, est souvent très important, ainsi que l'utilisation des tableaux ou des listes qui sont des dispositifs représentationnels permettant de restituer l'information sous une forme synthétique propice au traitement de l'information (Bertin 1973; Goody 1979). Toute cette information n'est, par contre, que très rarement indexée au sens où un algorithme ne pourra définir de manière déterministe les briques élémentaires de cette structuration.

<b>Rejets microbiologiques</b>	
<i>E-RM-1</i> : Atteinte du seuil réglementaire admissible de la concentration d'amibes dans l'eau de la rivière	<i>O-RM-1</i> : Traitement aux UV
<i>E-RM-2</i> : Renforcement de la réglementation de la concentration d'amibes dans l'eau de la rivière	<i>O-RM-2</i> : Changement des condenseurs en inox ou titane par des condenseurs mixtes
<i>E-RM-3</i> : Apparition d'un cas de méningo-encéphalite amibienne primitive malgré le respect des seuils réglementaires	<i>O-RM-4</i> : Traitement à la monochloramine <i>O-RM-5</i> : Chloration massive temporaire
<i>E-RM-4</i> : Prolifération dans la rivière d'amibes pathogènes malgré le respect des seuils réglementaires des rejets d'EDF	<i>O-Z-3</i> : Pré-traitement de l'eau d'appoint
<i>E-RM-5</i> : Non autorisation de traiter à la monochloramine pour les sites équipés actuellement de condenseurs en laiton qui vont être changés en condenseurs en inox ou titane	<i>O-RM-1</i> : Traitement aux UV <i>O-RM-5</i> : Chloration massive temporaire <i>O-Z-3</i> : Pré-traitement de l'eau d'appoint
<i>E-RM-6</i> : Découverte d'un nouveau traitement pour lutter contre la prolifération d'amibes pathogènes	<i>O-RM-3</i> : Industrialisation d'un nouveau traitement de lutte contre les amibes pathogènes
<i>E-RM-7</i> : Perte d'efficacité du traitement UV	<i>O-RM-10</i> : Amélioration du dispositif de traitement aux UV <i>O-RM-11</i> : Recherche d'un nouveau procédé non chimique
<i>E-RM-8</i> : Perte d'efficacité du traitement à la	<i>O-RM-11</i> : Recherche d'un nouveau procédé non chimique

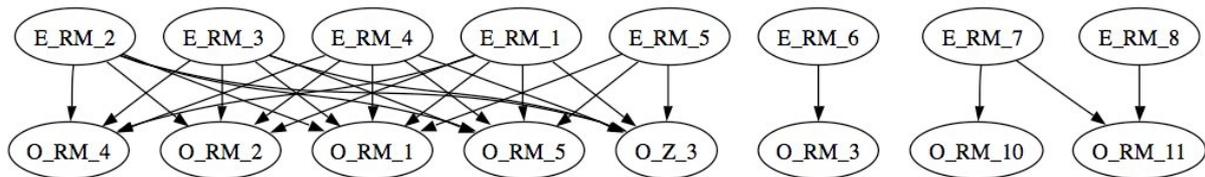


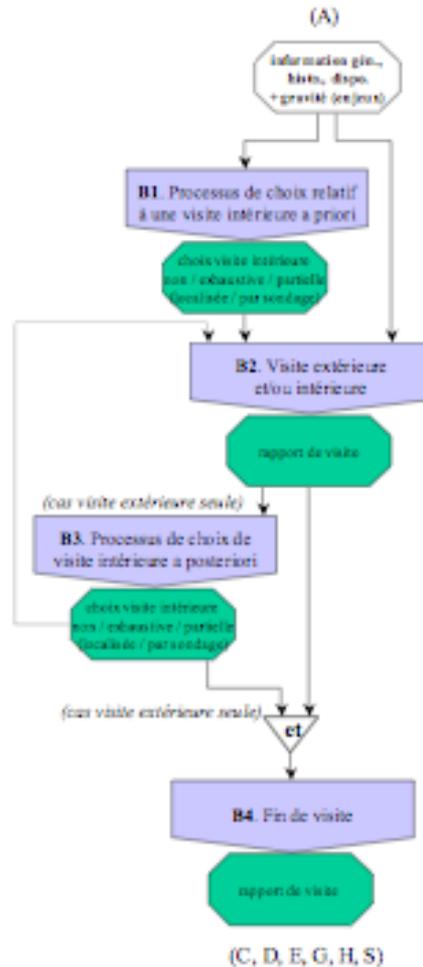
Figure 46 Exemple de tableau mis en regard avec la modélisation résultante

En effet, les auteurs ne précisent pas, sachant qu'il n'en ont pas nécessairement la possibilité, ce que représente sémantiquement la mise en forme exploitée (cf. Figure 46). On ne capte pas au niveau de la saisie l'intention de l'auteur (Nanard et Nanard 2005), ce qui représenterait pourtant une information essentielle, par exemple, dans la rédaction coopérative (Tazi et al. 2006). De même, la production de documents multimodaux, alliant texte et graphiques, par exemple, n'intègre pas nécessairement les liens existants entre fragments textuels et

représentations graphiques (cf. Figure 47). Au travers de l'exploitation de ce logiciel Wysiwyg, on a donc une dilution du contenu, au regard d'une exploitation par la machine, du fait de la fusion entre structuration et mise en forme. Cette dilution se développe au profit d'une facilité d'utilisation.

## 2.2 Phase B. Visite extérieure et / ou intérieure

### 2.2.1 Synoptique de la phase B



NB. L'étape B3 n'a lieu que dans le cas où la visite intérieure n'a pas été retenue en B1.

### 2.2.2 Détail de l'étape B1 – Processus de choix relatif à une visite intérieure a priori

Pour effectuer le choix de réaliser la visite intérieure ou non, on pourra se référer à la check-list **LE 4 « Choix d'une visite intérieure »** (☞ page 44) précisant les principaux éléments à prendre en compte.

### 2.2.3 Détail de l'étape B2 – Visite extérieure et/ou intérieure

On se référera aux check-lists **LE 5 « Éléments observables par visite extérieure tôle »** (☞ page 45) et **LE 6 « Éléments observables par visite intérieure tôle »** (☞ page 48) listant les éléments observables lors d'une visite. Ces listes d'éléments sont données à titre indicatifs, et ne sauraient être exhaustives.

Figure 47 Représentation multimodales issue du guide de diagnostic des conduites forcées

## 1.2. Excel au cœur d'une pratique de structuration de l'information

Le tableur Microsoft Excel est omniprésent dans les pratiques. Analystes, experts et informaticiens en font une utilisation massive et, même si chacun l'utilise à sa manière, il est possible de dégager quelques propriétés récurrentes dans son exploitation. Précisons tout d'abord que ce ne sont pas nécessairement les fonctionnalités de tableur qui sont, le plus souvent, valorisées. L'intérêt se porte en général sur son espace de représentation, sa surface d'écriture, ainsi que le mode de saisie qu'il introduit. L'espace de représentation se présente sous la forme d'un classeur composé d'autant de feuilles de calcul que l'on souhaite (espace quasiment infini de représentation). Chaque feuille de calcul représente un espace matriciel de saisie constitué de cellules. Les cellules peuvent être colorées, redimensionnées et fusionnées. L'observation de plusieurs documents produits par Excel, nous permet de dégager une pratique d'écriture reflétant une structuration hiérarchique des informations. Il s'agit d'un usage récurrent qui se traduit visuellement en termes de représentation par une fusion de cellule. Ainsi, une feuille de calcul telle que présentée sur la Figure 48, dont le contenu est majoritairement textuel, reflète une structuration hiérarchique de l'information. Sur cette figure, on a appliqué en surimpression une racine virtuelle qui est reliée aux différentes feuilles de calcul composants le classeur : c'est le premier niveau de hiérarchie. Le deuxième niveau de hiérarchie est défini par les différentes lignes. Le suivant est introduit lorsque les lignes se divisent et ainsi de suite. Au regard de la sémiologie graphique, une dimension du plan est associée à la profondeur de la hiérarchie et la seconde à la largeur. Le tableur Excel permet ensuite de particulariser les cellules par d'autres variables visuelles comme la couleur permettant de caractériser les entités élémentaires constituant la structure hiérarchique.

A travers ce mode de saisie, on identifie bien que l'usage qui est fait ici du tableur Excel est très différents de ce pourquoi il a été conçu. Ceci est tout à fait caractéristiques du passage de *l'outil à l'instrument* défini par Rabardel (Béguin et Rabardel 2000) lorsqu'un outil est intégré dans une activité et particulièrement une pratique utilisateur. En effet, un outil, un artefact, quel qu'il soit pourra avoir une utilisation très différente de ce pourquoi il a été conçu et ce, en fonction de l'improvisation, de l'astuce et de l'inventivité des utilisateurs. Dans la situation présente, le tableur Excel devient un moyen de modéliser et de structurer l'information, par une représentation matricielle sous un mode de saisie tabulaire. La granularité, l'entité élémentaire, de ce mode de structuration est alors la cellule : la feuille de calcul de Excel est alors un document très structuré.

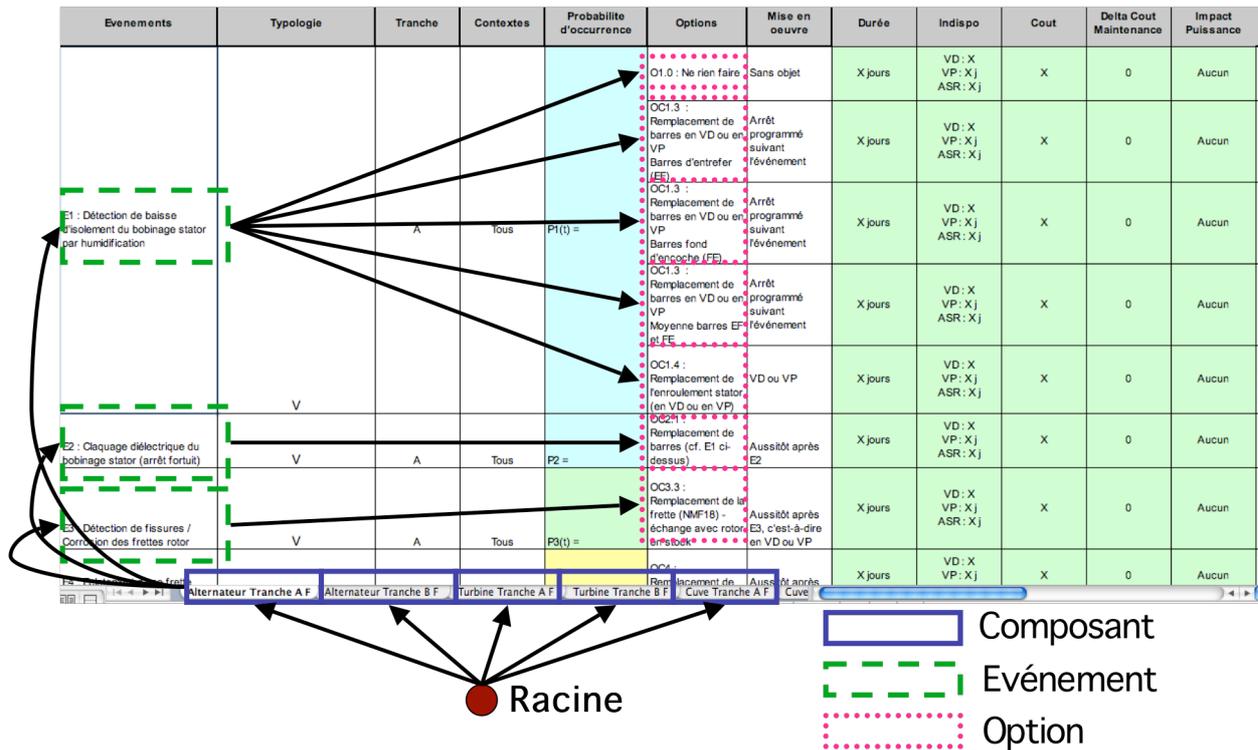


Figure 48 Fichier Excel exploité comme espace de structuration d'information

Chacun des utilisateurs d'Excel - qui dès lors prend un statut d'auteur - va à partir de cette brique élémentaire que constitue la cellule, ainsi que les capacités d'édition de l'architecte, dans le cas présent Excel, construire son propre langage de visualisation de l'information. La feuille de calcul du tableur constitue alors l'espace de modélisation pour l'analyste. Chaque analyste exploite alors la feuille de calcul qui propose un plan à 2 dimensions et des variables visuelles (couleur, taille...) au niveau des cellules, pour produire une image au sens de Bertin (Bertin 1973). Cette image tendra à mettre en relation la sémantique des informations avec les variables visuelles de Excel.

### 1.3. Du « WYSIWYG » au « WYSIWYM » : au-delà de la capitalisation du fond, capitaliser la forme

La pratique d'utilisation d'Excel que nous venons de présenter est bien entendu influencée par notre contexte d'étude. L'information capitalisée par les analystes est par nature très structurée du fait du milieu technique dans lequel nous évoluons. En plus du caractère assez technique des informations, les documents produits sont d'un type particulier. L'analyste produit des documents, par exemple, de capitalisation qui doivent expliciter, savoir et savoir faire et ce, de la manière la plus claire possible. Le niveau de discours exploité dans ces documents est alors assez « simple » - on exploite peu les subtilités du langage naturel

contrairement au récit par exemple - bien qu'intégrant un vocabulaire, une terminologie, propre aux différents métiers. On remarque que les capacités de mise en forme offertes par l'architexte sont exploitées de manière à bien véhiculer la sémantique de la modélisation produite par l'analyste. C'est le cas dans Excel où l'espace matriciel de saisie permet de représenter une information hiérarchique.

Bien évidemment chaque auteur introduit une part d'improvisation dans l'utilisation des architextes pour construire ses documents. Cependant, l'utilisation d'Excel comme outil de structuration hiérarchique de l'information est réellement un phénomène qui dépasse le cadre isolé de notre étude. Au-delà de cet exemple, les architextes, majoritairement Wysiwyg (What you see is what you get), offrent un espace d'édition infini et libre de représentation : l'exemple des feuilles de calcul Excel est caractéristique. Il est possible de combiner modalités textuelles et graphiques (diagrammes, schémas, images...) dans le document. Enfin, ces architextes offrent de larges possibilités de « copier – coller » que cela soit dans un document, ou au besoin entre différents documents produits par les différents logiciels.

Ainsi, l'analyste, ou plus généralement l'auteur d'un document, face à un architexte va définir son propre langage de représentation de l'information et ce, en fonction des possibilités de cet architexte. En intégrant l'architexte dans son activité, il le fait passer du statut d'outil à celui d'instrument (Béguin et Rabardel 2000) : il transforme l'architexte choisi en instrument de modélisation de connaissances. L'utilisation d'un architexte aussi commun que Word ou Excel, par exemple, permet à l'auteur, l'analyste, de conserver une autonomie dans la création du document (création du document qui correspond à la modélisation dans notre contexte de gestion des connaissances). Il peut définir son propre système sémiotique, en mettant en relation la sémantique des informations qu'il souhaite représenter avec les variables visuelles, mais également interactives, disponibles. C'est à partir de ce travail que l'auteur va définir le contrat de lecture et d'écriture dans le document. L'auteur modélise à travers la création de différentes instances de document. Au fil des versions, le système sémiotique, le langage, de représentation des informations va se stabiliser, permettant de dégager un modèle de document qui constitue, dans cette activité de gestion des connaissances, quasiment un modèle de connaissances. C'est bien le travail de construction des différentes instances, mêlant fond et forme, qui permet de définir à la fois la structure logique du document, mais également la sémantique qui va transiter par la visualisation (mise en forme matérielle) produite.

En définitive, l'analyste qui conduit la construction du document, va exploiter un architexte qui bien souvent correspond à un logiciel Wysiwyg puisqu'il construit une visualisation de

l'information telle qu'elle sera reproduite sur écran ou à l'impression. Ceci étant notre analyse nous amène au constat que l'analyste dans son activité définit des langages. Ces langages sont alors construites à partir des capacités de l'architexte en termes de visualisation (mise en forme), mais également d'interaction (liens hypertexte, affichage dynamique...).

Les architextes influencent la construction du document puisqu'ils possèdent des capacités en terme de définition de forme (la forme renvoie à la visualisation de l'information qui est produite), mais également des limites. Ainsi, ces architextes qui peuvent être des logiciels Wysiwyg (« What you see is what you get ») - concept associé au fait que la visualisation du document sur écran correspond à ce que l'on va obtenir, lors de sa diffusion électronique ou papier - vont parfois être exploités comme des logiciels Wysiwym (« What you see is what you mean »). Au travers des possibilités offertes (moyens graphiques et interactifs), les auteurs construisent leur propre langage de représentation de l'information qui va être plus ou moins explicité. Au travers de la définition de ces langages spécifiques, conçus par les auteurs, c'est le contrat de lecture et d'écriture qui est défini. Les auteurs possèdent alors une grande autonomie dans la construction des documents au regard du fond et de la forme.

## 2 - Analyse de la forme dans les outils d'aide à la gestion des actifs de production

Dans le cadre de la réalisation d'un document, la définition de la forme est du ressort de l'analyste. Il possède une certaine autonomie dans la définition des modalités de restitution de l'information. Lorsque l'on se place dans le contexte d'un système à base de connaissance, les interfaces d'accès aux informations sont généralement figées et ne permettent pas de recomposer le contenu que cela soit au niveau de son organisation ou encore de sa forme. Nous proposons de nous attarder dans un premier temps sur l'interface la plus couramment déployée : le formulaire.

### 2.1. Le formulaire

Le formulaire est souvent construit conformément aux structures de données implantées, ce mode de construction étant le plus évident dans la conception du système puisqu'il facilite la vérification de la cohérence entre le niveau applicatif et l'interface. Nous avons déjà discuté des limites de ce type d'approche. Elle tend majoritairement à faciliter l'accès aux structures de données et aux fonctionnalités du système sans nécessairement instrumenter la tâche de saisie.

Dans le cadre de l'Ingénierie des Connaissances, il est le plus souvent calqué sur le modèle de connaissances. Cette symétrie entre l'organisation du formulaire et le modèle induit que, sans connaissance du modèle, il peut être difficile d'exploiter l'outil proposé - cette symétrie n'étant pas forcément explicite sur la restitution de ce formulaire. Pour illustrer le caractère non explicite, nous pouvons par exemple évoquer dans la plupart des cas, l'absence de *représentation des relations entre les champs* constituant le formulaire ; chaque information paraît alors, en perception directe, canonique et indépendante ce qui ne se révèle pas propice à une appropriation des modèles en situation de saisie ou de consultation. Ainsi, en dehors de la consultation de documents externes à l'application, il n'est pas possible d'obtenir des informations expliquant tels ou tels champs de saisie ou encore leurs relations de dépendance. Les systèmes informatiques conduisent souvent à un ensemble de formulaires, qui permettent de visualiser la représentation des connaissances de manière fragmentée, sans pouvoir réellement en obtenir une vision synthétique et intégrée. L'accès se fait alors via cette succession de formulaires, qui engendre à chaque fois un changement de contexte pour l'utilisateur : ceci tend à créer des phénomènes de désorientation (Cox et Brna 1995) et à nuire à la compréhension. Dans ce contexte, reprendre une modélisation existante se révèle une tâche difficile, ou alors nécessitant un investissement intellectuel lourd, se fondant sur la capacité d'intérioriser les modèles pour, qu'à partir du formulaire, on puisse se construire une image de l'instanciation.

Le formulaire offre pourtant l'extrême avantage d'obliger les auteurs à se conformer à une structure formelle (Bachimont 2004). De ce fait, il ne représente pas un moyen d'écriture souple. Ce manque de souplesse est lié au fait, qu'il impose souvent la fourniture d'une information à un instant donné. Ceci est souvent dû au fait que le formulaire est en relation avec un système qui gère la cohérence des données de manière synchrone. Dans ce contexte, il n'est pas possible de laisser certaines informations indispensables, par exemple, au déroulement de l'algorithme incomplet et ce, même de manière temporaire. Ceci peut se poser en contradiction avec la modélisation, qui représente une activité créative dans laquelle il est parfois nécessaire de « laisser quelques blancs » pendant un moment. Le formulaire en tant que forme de restitution et avec le fonctionnement qu'il véhicule, peut alors contraindre à modéliser en dehors du système, pour ensuite saisir les informations au kilomètre : l'interface n'instrumente pas alors la tâche de modélisation, mais se contente de cadrer la saisie.

## 2.2. Des sources de données dissociées dans la restitution des informations

On s'intéresse ici à la restitution des résultats de sortie des algorithmes de simulation. Ils correspondent à des indicateurs statistiques et probabilistes, ainsi qu'à des traces textuelles d'exécution d'algorithmes. Ces traces textuelles décrivent la chronologie des événements et des options de certaines réalisations.

Les courbes semblent être le mode de représentation optimal. Cette adéquation entre indicateurs et courbes est relative aux conventions issues du domaine des statistiques, des probabilités et plus généralement des mathématiques. Leur pertinence n'est pas à mettre en doute, d'autant qu'il s'agit d'une restitution construite en tenant compte de la pratique des domaines cibles (Sûreté de fonctionnement, analyse de risques industriels). Toutefois, il semble que le lien entre les courbes et les données n'est pas nécessaire. Il n'existe pas de liaisons directes, manipulables, pour passer de l'image au texte dans une démarche de navigation, de lecture, où texte et image s'argumentent et se renforcent l'un avec l'autre : on dénote ici le manque de mise en contexte ou d'explication des informations proposées. Il faut alors passer par un post traitement qui se révèle « manuel ». Il se matérialise par la production d'un document qui « réalise » le lien entre résultats de sortie et données d'entrée par exemple.

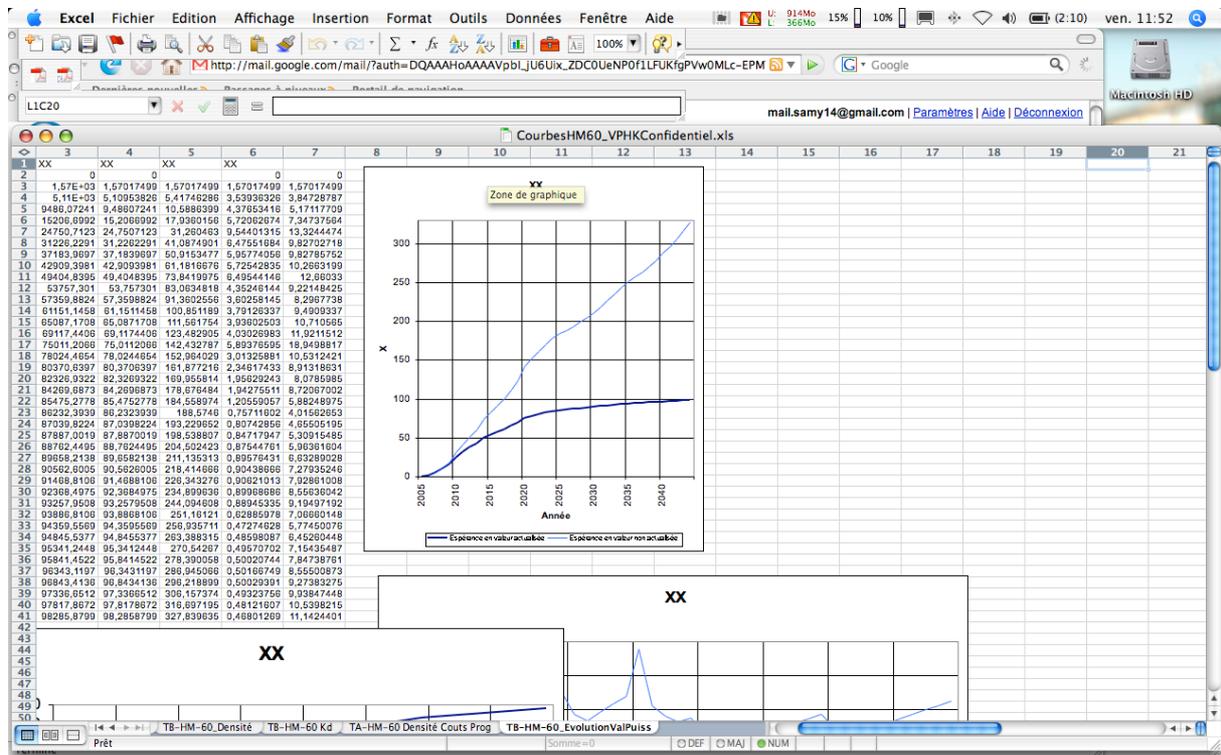


Figure 49 – Mise en forme et analyse de données issues du fichier des indicateurs au travers d'Excel

Là encore l'interface n'instrumente pas la tâche d'analyse des données, mais se contente de mettre à disposition le contenu. L'analyste, avant de présenter les résultats, va agréger un ensemble de données sous la forme d'un document en exploitant un architexte qui lui convient. Son expertise est essentielle dans cette tâche et il semble réellement difficile de pouvoir évincer l'analyste du dispositif (au profit de la machine par exemple). Il est, par contre, certainement possible de mieux instrumenter sa tâche.

### 3 - Activité de gestion de connaissances au regard de la dimension forme

L'analyse, sous l'angle de la dimension forme, illustre bien le rôle prépondérant du document. Le document, même s'il a évolué dans sa forme impactée par le support numérique, reste le vecteur privilégié de diffusion de l'information. L'activité de gestion de connaissance s'appuie sur la rédaction de nombreux documents, qui sont pour la plupart édités par des logiciels courants. Ces logiciels sont également choisis car ils peuvent être considérés comme les meilleurs éditeurs dans les systèmes de représentation qu'ils proposent.

L'analyste reste au cœur de la rédaction de ces documents et on assiste alors à un certain artisanat dans l'utilisation des logiciels. L'analyste profite d'une autonomie dans le choix de la forme, contrairement au cas du développement d'un système informatique, où la malléabilité de l'interface de diffusion d'information, par exemple, reste cantonnée essentiellement aux informaticiens. Les documents de capitalisation sont un cas intéressant puisqu'ils sont par nature très structurés. Comme nous avons pu l'observer, dans la rédaction de ces documents, l'analyste va exploiter un architexte et définir un langage de structuration, permettant de produire une représentation des connaissances. Ce langage de structuration va s'appuyer sur les moyens visuels et interactifs, disponibles au niveau de l'architexte. C'est par exemple la mise en forme matérielle du document qui permet de dissocier un certain nombre de briques élémentaires dans le document. L'analyste exploite alors des logiciels Wysiwyg en Wysiwym et se retrouve en position de concepteur d'un langage de représentation de connaissances, avec une marge de manœuvre contrainte par les possibilités de l'architexte choisi.

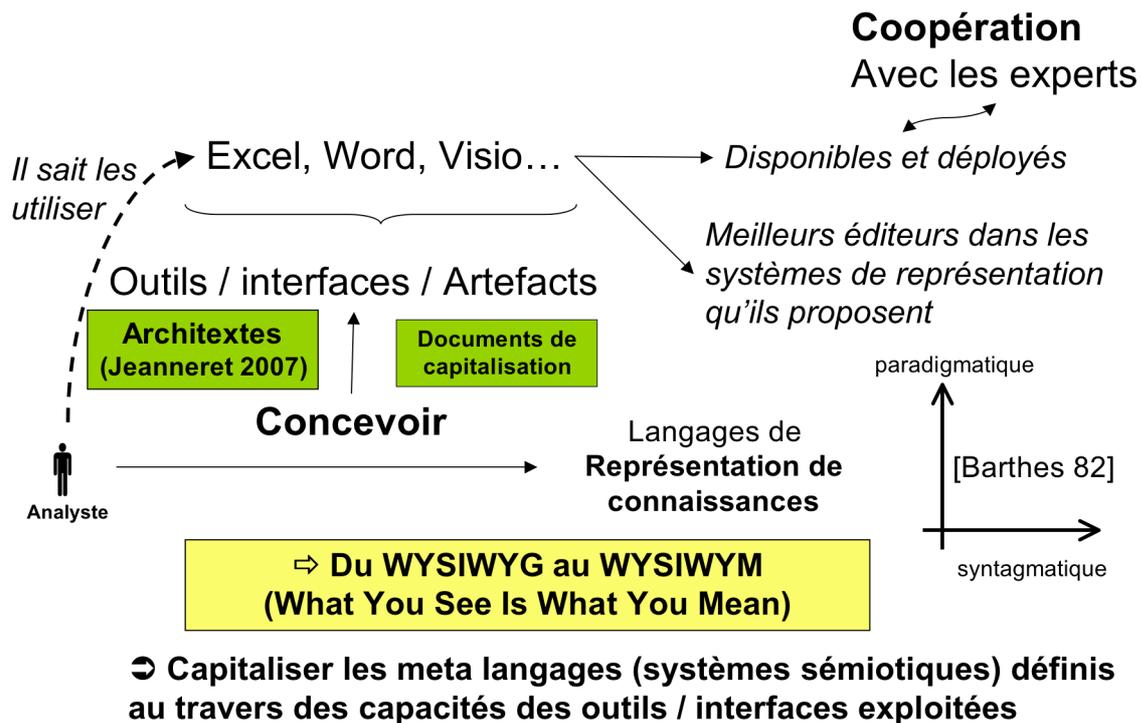


Figure 50 Le travail de l'analyste : concevoir des langages

Ce mode d'utilisation des architextes par les analystes pose alors le problème d'une dilution possible de la sémantique de la capitalisation produite, si le langage ainsi défini n'est pas identifiable « naturellement » ou explicité. En effet, les auteurs n'explicitent pas nécessairement ce que représente sémantiquement la mise en forme exploitée. Au niveau de la saisie, il n'y a pas de capture de l'intention de l'auteur (Nanard et Nanard 2005), sous-jacente à la mise en forme définie, sachant qu'il n'en a pas toujours la possibilité. Ceci représenterait pourtant une information essentielle par exemple dans la rédaction coopérative (Tazi et al. 2006). Il s'agit également d'un point essentiel dans la pérennisation des informations capitalisées (conservation du contrat de lecture et d'écriture). De plus, étant donné le volume de documents produits aujourd'hui, il est important que ces derniers puissent être analysés par la machine, et notamment dans l'activité de gestion des connaissances, que l'ordinateur soit capable d'identifier les briques élémentaires qui constituent une modélisation produite sous la forme d'un document.

Un autre cas de figure est la mise en œuvre de systèmes à base de connaissances s'appuyant sur une représentation formelle de connaissances sur laquelle l'ordinateur peut raisonner. Le formulaire est l'interface la plus courante pour la saisie de donnée, car il permet d'obliger les auteurs à se conformer aux structures de données et à une saisie procédurale : il en résulte une saisie très structurée des informations. La visualisation de l'information est pour une part très

claire, puisque le découpage en champs de saisie spécifie bien les informations à fournir, mais il pâtit également des contraintes qu'il introduit dans la saisie. En outre, il n'offre pas nécessairement une vue très globalisante. Il peut être nécessaire de passer d'un formulaire à un autre, pour identifier les liens entre plusieurs informations, et ainsi parvenir à ne construire une vision globale.

Dans le cadre d'un système à base de connaissances, nous avons également évoqué la problématique de la restitution des résultats de sortie. L'accent a été mis sur le caractère très figé des interfaces proposées. Les informations sont ainsi rendues, disponibles sans réelle malléabilité sur leur forme. De nouveau, c'est le document qui est l'alternative aux interfaces informatiques. Il est rédigé par l'analyste qui va mettre en relation résultats d'entrée et de sortie. La rédaction de ce document passe par un architecte qui va permettre de mixer des sources de données diverses. La réalisation est très largement emprunte de l'expertise de l'analyste qui va à la fois interpréter les résultats produits et les argumenter. Il pourra également dans un souci de lisibilité changer la mise en forme des sources de données exploitées.

## Chapitre 9 - Conclusion d'analyse

### 1 - Caractérisation de l'activité de gestion des connaissances

Nous venons de décrire une analyse de l'activité de gestion des connaissances selon les trois dimensions medium, signe et forme. Cette analyse empirique nous a permis de dégager un ensemble d'éléments caractéristiques en relation avec les démarches de gestion de connaissances.

Tout d'abord, nous mettons en avant le caractère incontournable de la forme documentaire comme support à la capitalisation des connaissances et à la diffusion d'une représentation. Que la vocation du projet soit de produire effectivement un document ou qu'il s'agisse de la mise en œuvre d'un système à base de connaissances, le document reste la forme privilégiée du processus de restitution d'information et, bien souvent, le plus adapté dans la plupart des situations. Le document constitue alors le medium dans la médiatisation des informations au sein du collectif engagé dans l'action de gestion des connaissances.

#### 1.1. Un registre d'écriture particulier pour les documents de capitalisation

Les documents produits, quels qu'ils soient, se caractérisent par un registre d'écriture particulier. Le contenu en est très structuré dans l'expression, la structuration étant également renforcée ou définie par une mise en forme importante et l'utilisation de nombreux dispositifs d'organisation d'informations tels que les listes indentées, les tableaux, les diagrammes, etc... Cette structuration très importante tient à la fois au caractère technique des informations et au milieu dans lequel nous évoluons. Elle est également liée à l'activité de capitalisation, elle-même qui doit produire une indexation très fine des contenus. L'indexation correspond à une identification des briques élémentaires d'information et à leur caractérisation : il s'agit par exemple d'être en mesure de délimiter dans un document des fragments pour en préciser la nature. Dans cette activité d'indexation, il convient de proposer un environnement de conception de document permettant d'instrumenter au mieux cette structuration.

#### 1.2. Une diversité d'outils mobilisés dans la création des documents

L'observation de la production de ces documents de capitalisation a permis d'identifier un large éventail d'outils employés. Au-delà de cet éclectisme, les analystes exploitent ces outils

avec un certain artisanat. Ils développent leurs propres stratégies d'utilisation pour transformer l'outil en réel instrument support à leur activité (Béguin et Rabardel 2000). Ainsi, traitement de texte ou tableur sont employés comme atelier de modélisation et de structuration d'information. Les analystes s'appuient sur ces logiciels tout d'abord parce qu'il savent les utiliser ; ces logiciels faciliteront la coopération avec les experts par exemple dans l'activité de conception. Ensuite, ils aident à la diffusion d'une restitution de connaissances. Enfin, et c'est certainement un des point les plus importants, l'utilisation d'un logiciel courant permet à l'analyste de conserver une grande autonomie dans la définition de la forme de restitution. Ceci ne serait pas le cas confronté au développement d'une interface informatique qui nécessiterait de faire appel à un informaticien pour la modifier. Le document offre ainsi une malléabilité et une flexibilité sur la forme et le fond, ce qui accentue l'intérêt de la restitution documentaire et souligne l'importance de s'appuyer sur les infrastructures logicielles déployées et les logiciels auxquels nos acteurs sont habitués.

### 1.3. L'importance de l'autonomie de l'analyste

L'autonomie de l'analyste lui permet de définir un langage de structuration en situation de capitalisation. Plutôt que de devoir anticiper le modèle de connaissance, il est conçu de concert avec la capitalisation des connaissances elles-mêmes. Le langage de structuration, qui est alors défini, tient tout aussi bien à des éléments textuels insérés dans le texte qu'à des éléments de mise en forme qui vont délimiter des fragments du document. Cette structuration du contenu s'appuie donc et va être contrainte par les possibilités de l'outil et de l'architexte, choisis par l'analyste. Dans cette situation, le langage, défini par l'analyste, n'est pas nécessairement sous une forme perceptible par la machine. Ainsi, au travers de l'exploitation de ce logiciel Wysiwyg, même s'ils sont exploités en Wysiwym, on a une dilution de la sémantique du document, au regard d'une exploitation par la machine due à un manque d'indexation (au sens identifiable par la machine) des systèmes sémiotiques mis en œuvre. Un humain pourra aisément discerner la structure visuellement. Il est envisageable de mettre en œuvre une indexation en texte intégrale, ou encore des analyses de text-mining pour dégager une indexation « pertinente ». Cependant, ces mécanismes automatiques ou semi-supervisés ne seront jamais aussi précises qu'une indexation manuelle. Elles répondent au traitement des vastes corpus de documents existants. Nous ne remettons néanmoins pas en cause ces approches, mais avec l'accélération produite par le numérique, il nous semble important de mettre à disposition des utilisateurs, les outils permettant de construire cette indexation à la création des documents, plutôt que d'analyser a posteriori leurs structurations.

#### 1.4. Rédaction d'informations relationnelles

La structuration du contenu des documents produits ne se limite pas à un découpage de fragments documentaires. La dimension structurelle se concrétise également par la mise en relation d'informations : c'est notamment le cas dans les hypertextes. La construction de cette information relationnelle est accessible via des outils d'édition de documents courants lorsque le contenu est restreint. Cette tâche se complexifie lorsque le contenu s'étend car il peut alors être difficile d'avoir une vue globale de son ensemble. En effet, le manque de vue d'ensemble ne permet plus de bien identifier les briques élémentaires que l'on souhaite mettre en relation. La saisie devient alors complexe, pouvant aboutir à développer de nombreuses incohérences dans la rédaction. Ce problème de mise en relation d'informations se fait également ressentir dans le développement des systèmes à base de connaissances.

#### 1.5. Une dichotomie entre système à base de connaissances formel et les documents

Dans ce cas de figure, les documents, comme nous avons pu l'évoquer, restent l'interface privilégiée en amont pour la capitalisation et en aval pour les systèmes avec une restitution du contenu des bases de connaissances ou des résultats de sortie. Ces documents sont alors le plus souvent indépendants du système. Il n'y a pas de liaison entre leur contenu, bien qu'il soit numérisé, et les informations provenant du système informatique lui-même. Cette situation tend à développer des décalages et parfois des incohérences entre les documents et les informations contenus dans le système. Ce point pose également le problème de multiplier la saisie de nombreuses informations qui vont devoir être à la fois reportées dans les documents et saisies dans les systèmes informatiques.

## 2 - Le document structuré comme support à la capitalisation et la restitution

En réponse au point que nous avons évoqué à l'issue du chapitre précédent, nous proposons une approche à la fois méthodologique et instrumentale basée sur la mise en œuvre de document numérique structuré. Le document numérique structuré dans notre approche doit alors devenir un support à l'activité de gestion de connaissances qui va tout à la fois supporter l'activité de capitalisation et constituer une solution flexible de restitution d'information.

## 2.1. Apports du document numérique structuré dans la capitalisation

Comme nous l'avons déjà évoqué dans ce rapport, l'activité de capitalisation des connaissances représente une tâche peu déterministe. Il est en effet difficile d'anticiper les tenants et les aboutissants de ce type de démarche. Nous proposons d'illustrer les apports du document structuré face à deux situations : la rédaction d'un document et la réalisation d'un système informatique. Dans le cadre de la rédaction d'un document de capitalisation, nous aborderons l'apport face à :

- La dilution du contenu

Dans la réalisation d'un système à base de connaissances, il s'agit d'illustrer l'apport dans :

- La modélisation
- La saisie d'informations dans un cadre formel

### *2.1.a La dilution du contenu*

L'utilisation de documents numériques structurés dans le dispositif de capitalisation doit permettre, au moins pour partie, de pallier les inconvénients de l'informatique bureautique tels que présentés précédemment. Nous avons bien souligné que l'utilisation de ces outils offrait une grande autonomie pour les analystes sur la définition de la forme que prendra la capitalisation : c'est essentiellement le cas dans la production des documents.

Cette production de documents de capitalisation au travers des outils de bureautique centrés sur le concept de Wysiwyg tend à produire des documents très structurés et très codifiés (cf. Chapitre 8 -1 - page 148). La structuration et la codification, si elle est appréhendable visuellement, n'est pas nécessairement identifiable par la machine. La valorisation des documents et en particulier de leur contenu est alors rendu difficile ; il peut être complexe de dissocier les briques élémentaires d'informations qui y sont contenues, ainsi que leur mise en relation dans le document.

Par l'utilisation de documents structurés, il s'agit d'intégrer, dès la rédaction des documents, des éléments structuraux qui permettront d'identifier et de caractériser des fragments de documents. Les éléments structuraux pourront tout aussi bien correspondre à un balisage logique ou généraliste (cf. Chapitre 3 -1.3.b page 67) : ils représentent alors une indexation fine du document. Cette indexation sera alors identifiable par l'ordinateur permettant d'opérer des traitements sur le document.

### *2.1.b Dans la modélisation*

Dans l'activité de modélisation des connaissances, il peut être difficile d'anticiper la définition du modèle qui constituera par exemple le langage de représentation des connaissances. Bien que certains objectifs puissent être définis initialement, le contexte multi-domaines et multi-points de vue, ainsi que la transcription de connaissance, oblige les experts à réinterroger leur savoir ; cela induit qu'il est souvent difficile de réellement prévoir les finalités de ce genre de démarche. Les démarches de conception trop prédictives pâtissent alors du fait de vouloir spécifier des modèles trop tôt et trop en amont. Les modèles se construisent souvent à mesure de l'activité de capitalisation elle-même imposant de travailler sur des cycles courts de conception.

La mise en œuvre de documents structurés doit permettre de s'écarter d'une approche de capitalisation basée sur une anticipation des structures de données. Le document structuré va plutôt constituer un espace de saisie uniquement cadré par le fait de produire un document bien formé. Il sera alors possible d'y inscrire des informations de manière « constructiviste » tout en les délimitant par des éléments structuraux qui identifieront les entités élémentaires du document. Ainsi, l'utilisation du document structuré doit se faire ici sans volonté de vouloir définir un modèle de document a priori, mais bien par la création d'instances de document. Un exemple de ce type d'approche en développement logiciel est l'extreme programming (Casabianca 2005)).

### *2.1.c Dans la saisie d'informations formelles*

La mise en œuvre de structures de données formelles exploitables par une logique de calcul implantée amène le plus souvent à produire des interfaces de modélisation trop cadrées qui tendent à une utilisation « taylorisée ». Un exemple caractéristique correspond à la mise en œuvre de systèmes informatiques basés sur un éventail de formulaires. Le formulaire constitue sans doute, actuellement, le moyen le plus efficace « *pour encadrer les auteurs* » et les pousser à se conformer à un langage formel (Bachimont 2004). Il est particulièrement pertinent pour garantir la cohérence des informations saisies, mais constitue également une contrainte dans le processus d'écriture. Le formulaire, et plus largement une succession de formulaires, impose une démarche de saisie (cf. Figure 51) (cf. Chapitre 8 -2.1. ) pour une analyse du formulaire. La multiplicité des formulaires peut également faire naître des phénomènes de désorientation chez l'utilisateur rendant difficile la construction d'une vision globale sur l'espace informationnel.

Le document structuré, sans volonté de le remplacer, peut alors se poser en alternative de ce type d'interface homme information. Il peut, en effet, être mis en œuvre de manière à proposer un cadre de saisie très formel tout en offrant une flexibilité, une souplesse, dans le processus d'écriture. Le cadrage formel est alors spécifié par un modèle de document qui traduira les contraintes formelles d'écriture des données. Les contraintes formelles correspondront à des contraintes lexicales et syntaxiques spécifiées dans le modèle de document structuré. Dans la rédaction du document, il sera possible d'écrire les informations selon sa convenance en terme d'ordre par exemple, voir même de laisser de manière temporaire des informations incomplètes (cf. Figure 51). Il s'agira seulement à l'issue de la rédaction du document que celui-ci soit en accord au modèle de document spécifié : le document structuré doit être valide. Notons que le modèle de document qui spécifie à la fois les contraintes syntaxiques et lexicales du document, va permettre de fournir une aide à la saisie avec un éditeur adapté proposant des mécanismes de complétion (cf. Figure 53). Ce modèle de document sera également en mesure de fournir une explication sur la nature des éléments structuraux proposés s'il est documenté (Marcoux 2006).

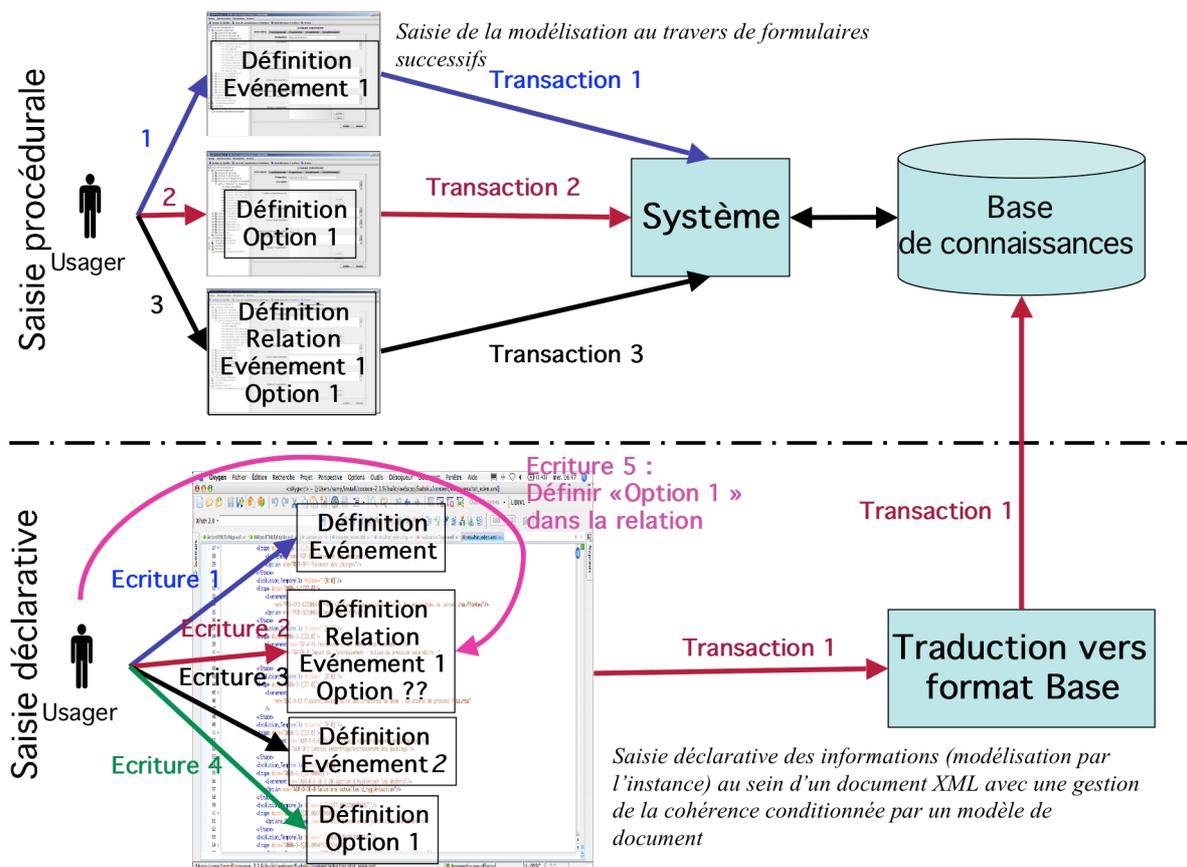


Figure 51 La saisie procédurale au travers d'un formulaire au regard d'une saisie déclarative dans un document structuré

```

<Evenements>
  <Evenement>
    <Libelle></Libelle>
    <Description></Description>
    <Famille></Famille>
    <Gravite></Gravite>
    <Repercution></Repercution>
    <Incontournable></Incontournable>
    <Unicite></Unicite>
    <ContexteAS>
      <Libelle></Libelle>
    </ContexteAS>
    <Informations_loi></Informations_loi>
    <Influence></Influence>
    <Options></Options>
  </Evenement>

```

Figure 52 L'éditeur de code source souligne en rouge les parties du document non valides

```

<Historique></Historique>
<Surveillance></Surveillance>
<Maintenance></Maintenance>
<Evenements>
  <Evenement>
    <Libelle></Libelle>
    <Informations></Informations>
    <Comportement></Comportement>
    <Classe></Classe>
    <Famille></Famille>

```

Figure 53 L'éditeur de code source, en s'appuyant sur le modèle de document propose les balises valides qui peuvent être introduites

## 2.2. Apports du document numérique structuré dans la restitution

Le choix du document numérique structuré comme interface de restitution est guidé tout d'abord par le fait qu'il introduit une réelle souplesse dans la visualisation des données. Il constitue une abstraction du document dissociant les données de leur mise en forme. Ceci permet, de manière aisée, de passer d'un mode de visualisation à un autre, via la définition de feuille de style par exemple. Il est ainsi possible de produire des documents textuels avec une mise en forme matérielle complexe, des schémas, des images... Il est également possible de combiner plusieurs types de visualisation à partir d'un unique document structuré. On aboutit alors à la production de documents multimodaux, intégrant des références intermodales facilitant la lecture (Bestgen et Dupont 2003) entre les différents types de visualisation. La dissociation entre données et mise en forme apporte des éléments de réponse à la dilution de la sémantique du document, tels qu'évoqués précédemment. Elle permet également d'avoir

une inscription de la sémantique visuelle mise en œuvre, une feuille de style pouvant constituer la légende si cette dernière est documentée.

Au-delà de la forme, les entités élémentaires qui constituent le document structuré devenant manipulables par l'ordinateur, de nombreux traitements plus ou moins complexes sont envisageables sur le contenu. Ces traitements peuvent opérer un filtrage ou une réorganisation du contenu. Ils peuvent également mettre en relation différents documents ou plus largement différentes sources d'information : ceci peut aboutir à la construction assistée ou dynamique d'hypertextes.

Le document structuré peut être manipulé de manière aisée, ouvrant de nombreuses perspectives de transformation et d'adaptation. En particulier, le choix d'un document structuré basé sur un standard ouvert peut permettre de l'adapter à de nombreux architextes existants. La condition nécessaire reste que ces architextes proposent également un format ouvert.

## Partie IV - Proposition : une approche centrée sur le concept de document

# Chapitre 10 - Approche centrée sur le document structuré

Notre approche propose d'exploiter le document structuré comme support à l'activité de gestion des connaissances. Nous avons introduit le document structuré au Chapitre 3 -1.3. page 65. A l'issue du chapitre précédent, nous avons décrit nos motivations quant à l'utilisation du document structuré comme support à la capitalisation et la restitution dans le cadre de démarches de gestion des connaissances. Ce chapitre propose de décrire avec précision l'utilisation et l'impact du document structuré au niveau de la capitalisation puis de la restitution.

## 1 - Document structuré et capitalisation : l'écriture

L'activité de capitalisation qui correspond à une tâche d'identification et de structuration d'information, constitue un travail très peu déterministe. Il est alors difficile d'anticiper sur la définition du modèle de structuration en la décorrélant de l'activité de capitalisation elle-même. Nous proposons donc que le document structuré devienne le support d'inscription de la capitalisation. Il pourra constituer un espace de conception de modélisation. L'écriture dans le document pourra tout d'abord être contrainte par le respect de règles rudimentaires relatives au fait de produire un document bien formé.

### 1.1. Le document structuré comme espace de conception du langage

Dans la méthode de capitalisation que nous proposons, l'activité de modélisation va correspondre à l'écriture d'un document bien formé. Nous rappellerons qu'un document structuré bien formé est un document qui répond aux règles lexicales et syntaxiques élémentaire du type de document structuré choisi (cf. Chapitre 3 -1.3. page 65).

Au travers de l'écriture de documents bien formés, il s'agit de fournir un espace de modélisation qui permette de faire évoluer le modèle de structuration et le contenu informationnel de concert. En effet, l'écriture se fait alors « au kilomètre ». Se faisant, on enrichit le contenu du document tout en spécifiant des éléments structuraux qui vont délimiter ce contenu.

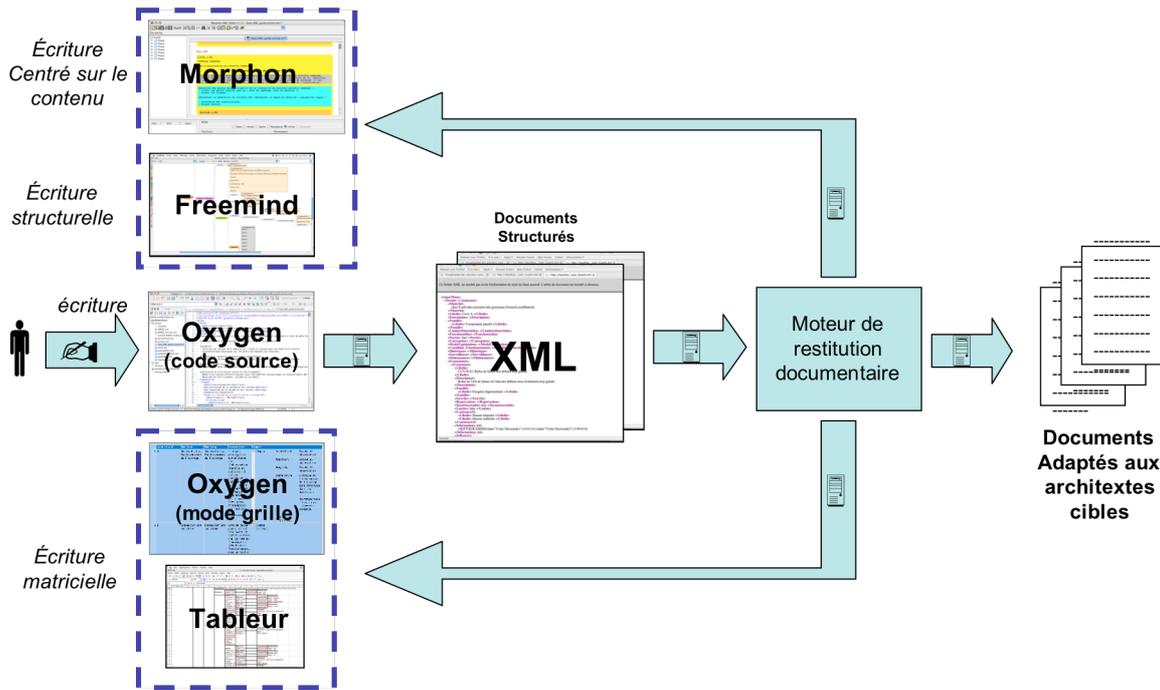


Figure 54 La capitalisation au travers de documents structurés

### 1.1.a Capitalisation et indexation d'entités élémentaires de contenu

Les éléments structuraux vont alors tout aussi bien traduire un balisage pouvant être appréhendé comme logique ou généraliste. La structuration de contenus permettra de faire transparaître des structures « entité-relation ». Les éléments structuraux insérés préciseront la nature des concepts manipulés produisant une indexation fine du contenu. Si l'on s'appuie sur le standard XML, comme solution technique de développement de document structuré, c'est la hiérarchie inhérente à l'enchevêtrement des balises qui va traduire la structure « entité-relation ». La hiérarchie se matérialise par l'indentation des informations sur la Figure 55. Les éléments structuraux, correspondant aux balises, pourront alors être dissociés, selon qu'ils contiennent un contenu textuel ou non, et selon qu'ils contiennent des éléments structuraux ou non. En fonction de ces critères, l'élément structurant n'est pas de même nature ou ne traduit pas un contenu de même nature.

Notons que si un modèle de connaissance préexiste, la définition d'un modèle de document permettra de transformer le document en espace de saisie très cadré, mais offrant une certaine souplesse dans l'activité d'écriture des données. Le document structuré restera alors le support d'inscription à la capitalisation, mais le langage de structuration sera cadré et contraint par le modèle de document.

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <!DOCTYPE Guide SYSTEM "guideEditionEtape.dtd">
3 <Guide>
4   <Phase>
5     <Identifiant>D</Identifiant>
6     <Nom_Court>Observation du fonctionnement de l'ouvrage</Nom_Court>
7     <Nom_long>Observation du fonctionnement de l'ouvrage</Nom_long>
8     <Commentaire>Le moyen principal de surveillance est l'observation visuelle et auditive. Il est par conséquent indispensable de lui donner
9 Cette phase d'inspection est réalisée par l'exploitant
10 </Commentaire>
11   <Etapes>
12     <Identifiant>Écoute et observation</Identifiant>
13     <Nom_Court>Écoute et observation </Nom_Court>
14     <Nom_long>Écoute et observation</Nom_long>
15     <Commentaire>L'écoute et l'observation de l'ouvrage sans démontage dans les domaines :
16 - Hydromécanique
17 - Génie civil
18 - Contrôle commande
19 </Commentaire>
20   <Etape Nom="Observation du fonctionnement hydro-méca ">
21     <Identifiant>D.1 - HM</Identifiant>
22     <Fiche_de_visite>
23       <Identifiant>D.1.A - HM</Identifiant>
24       <Domaine>HYDRO-MECANIQUE</Domaine>
25       <Nom_Court>Fonctionnement hydro-méca de l'écluse</Nom_Court>
26       <Nom_long>Fonctionnement hydro-méca de l'écluse</Nom_long>
27     <Commentaire>L'écoute du fonctionnement sans démontage :
28 - Placé à proximité des organes en mouvement (système de manoeuvre des portes, des vannes aqueduc, accouplement, pallier), il s'agit d'écouter
29
30 </Commentaire>

```

Figure 55 Exemple d'un fichier XML

### 1.1.b Le document structuré pour définir un langage

Dans la méthode de capitalisation que nous proposons, le modèle de structuration d'informations correspond au modèle de document en cours d'écriture. Cette démarche permet d'éviter à avoir à anticiper sur le langage de structuration de l'information. Elle permet également de construire celui-ci au fur et à mesure de la capitalisation : le modèle se dégage au travers de la construction d'une instance de document structuré.

Le fait de travailler sans définir de modèle de document a priori offre une grande autonomie à l'analyste dans la création de son langage de structuration. Le modèle de document qui constitue alors notre modèle de représentation de connaissances résulte de la création d'une instance de document structuré. A partir d'une instance, il est possible de dégager, ou même d'extraire automatiquement, le modèle de document correspondant (cf. Figure 56).

Le document structuré est donc exploité comme un réel espace de conception. Le document produit n'a pas vocation à être directement une version cohérente et optimale. Il est un réel instrument de capitalisation et de structuration de l'information. Il constitue ainsi un réel Document pour l'Action passant d'une version à une autre. Ce statut de DopA que prend ainsi le document est également lié au fait qu'il constitue un réel objet de coopération. Le document structuré allie une souplesse de structuration et de définition du langage de structuration, pour répondre aux besoins de l'analyste ; il permet également une indexation formelle (au sens informatique) des entités élémentaires assez importantes pour être retravaillées et exploitées par l'informaticien pour une valorisation plus complexe du contenu.

• Démarche de capitalisation

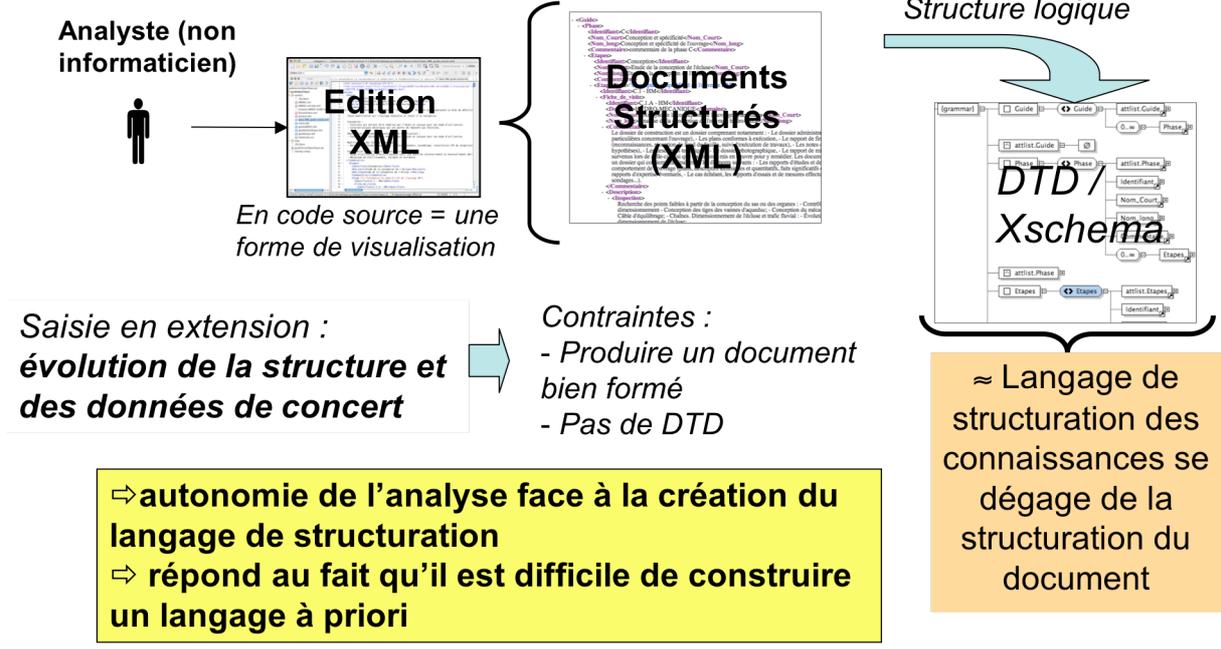


Figure 56 Utilisation du document structuré comme support à l'activité de capitalisation des connaissances

Dans la mise en œuvre de l'approche proposée, notre choix s'est porté vers le standard XML pour développer les documents structurés. L'utilisation de XML dans le domaine de la gestion des connaissances n'est pas à démontrer, puisqu'elle a déjà été largement discutée dans les domaines de la structuration ou de la restitution d'informations. La séparation fond/forme représente un atout souvent mis en avant qui permet de produire, à partir d'une même structure, différents modes de restitution (Dieng-Kuntz et al. 2001). Notre approche s'appuie sur ce standard, ce qui, en outre, nous permet de bénéficier de l'offre logicielle très riche qui l'accompagne. Nous proposons ici de décrire la mise en œuvre de notre approche en mettant l'accent sur les outils ayant permis de produire la capitalisation. Les éléments de retour d'expérience font essentiellement référence au projet relatif au développement des guides de diagnostic (cf. Chapitre 1 -3.2.a page 43).

1.2. Le document structuré comme support et mode d'écriture dans la capitalisation : retour d'expérience

L'expérience que nous avons mise en place a consisté à confronter un analyste non informaticien à un éditeur de code source XML. Le choix de l'éditeur s'est porté vers Oxygen (<http://www.oxygenxml.com/>). Il offre un espace classique d'édition, c'est-à-dire un espace

infini d'écriture adapté à l'écriture de code XML, accompagné de fonctionnalités classiques, tel que le « copier-coller ». Certaines fonctionnalités sont plus spécifiques aux langages balisés, puisque l'éditeur propose des mécanismes de complétion du balisage permettant d'accompagner l'activité de saisie. Il s'agira, par exemple, du support contextuel à la fermeture de balise. L'éditeur opère également une vérification permanente, a minima, du caractère bien formé du document en cours de réalisation. Dans la même lignée, si un modèle de document est spécifié, il vérifie le caractère valide du document *in vivo*, mettant en valeur les éventuelles erreurs ou incohérences.

### *1.2.a Document structuré et formalisation par un non informaticien*

La prise en main de l'éditeur et l'apprentissage des principes du langage structuré se sont faits assez rapidement par l'analyste. La possibilité de définir son propre langage de structuration a représenté une réelle souplesse dans la démarche de capitalisation. L'analyste a ainsi apprécié cette manière de structurer ses contenus. Il évoque même que « *l'écriture au travers de l'éditeur de code source XML n'était pas très différente de l'écriture dans un traitement de texte* ». La formalisation des connaissances se fait alors en extension, c'est-à-dire de manière constructiviste, en créant les balises nécessaires au gré des besoins. La visualisation pour l'auteur, l'analyste, correspond alors à un code source colorisé, et dont l'indentation traduit l'arborescence. Structure et contenu sont construits et évoluent en parallèle. Le document structuré prend alors une dimension d'espace de conception jusqu'à ce qu'il se stabilise. Cette stabilisation du document permet alors d'identifier le modèle de document – ceci se limite à exploiter les fonctionnalités de l'éditeur, qui permettent de calculer la DTD ou le Xschema du document courant. Ce modèle correspond à sa structure canonique, structure à partir de laquelle il est possible de dériver un ensemble de documents appartenant à une classe homogène. Il répertorie l'ensemble des éléments structuraux, les balises, qui constituent le lexique utilisable pour la structuration des informations. Il intègre également les règles d'imbrication des éléments structuraux spécifiant ainsi les règles syntaxiques de structuration.

Cette démarche et cette instrumentation de la capitalisation n'est toutefois pas sans poser problème. Il est bien évident qu'une formalisation déclarative en extension risque d'introduire des difficultés au niveau de la cohérence et être à l'origine de problèmes de « duplication » du contenu. On ne répond plus à la propriété d'unicité des données chères aux informaticiens. C'est, en effet, un phénomène que nous observons dans nos expérimentations. Ceci ne constitue toutefois pas, de notre point de vue, un écueil. En effet, nous pensons que ce travail

préliminaire de l'analyste - qui débouche sur un tel contenu - est essentiel, car il est toujours difficile d'avoir, a priori, une vue d'ensemble sur les éventuelles redondances que l'on aura à traiter. Dans la rédaction d'un tel document, l'analyste introduit un point de vue relatif à l'organisation de l'information qui l'aide dans son activité et sa réflexion : il crée son propre langage de structuration. Par contre, l'intégralité du contenu étant sous une forme très structurée, ces problèmes pourront être traités, de manière plus ou moins complexe, d'abord de manière semi-supervisée, ensuite en collaboration avec l'informaticien. Ainsi, les duplications de contenu pourront faire l'objet d'une factorisation, par exemple, dans un document indépendant.

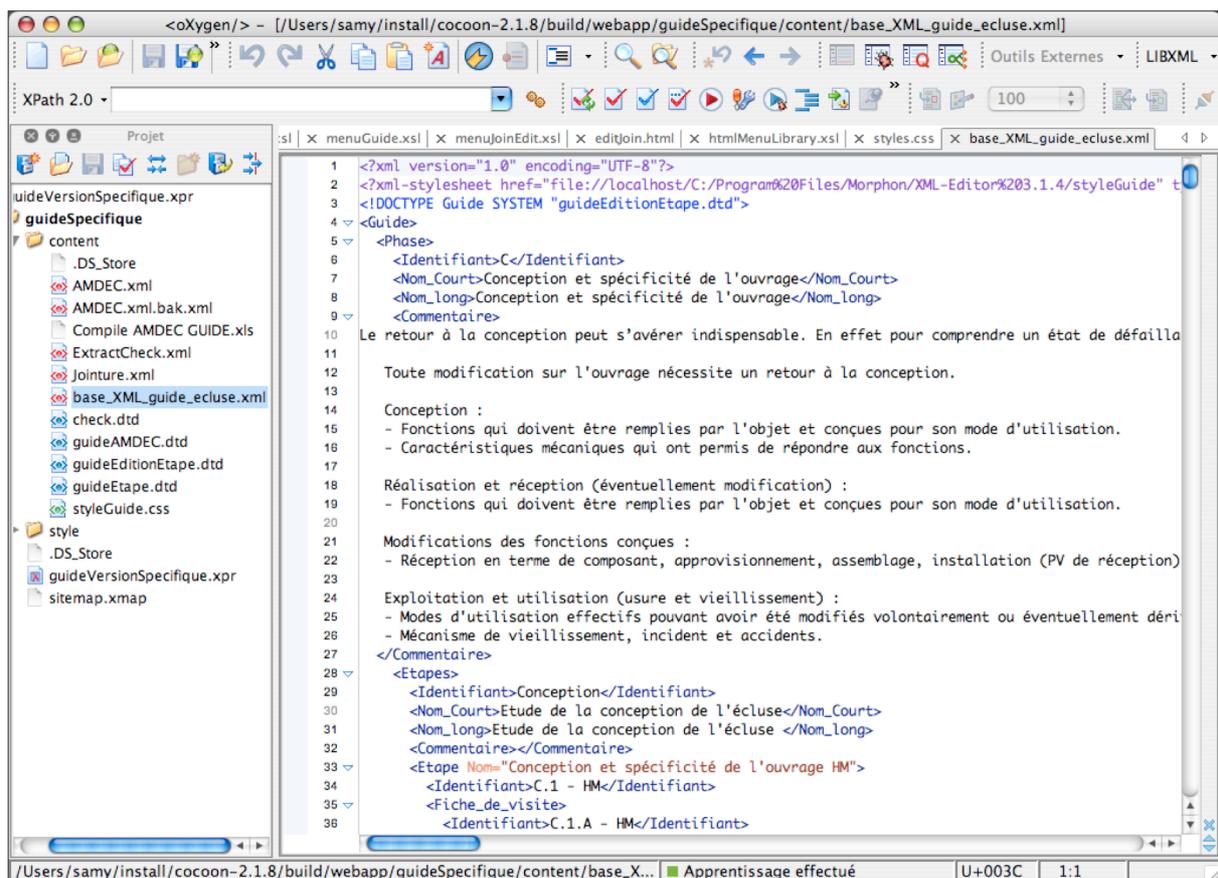


Figure 57 L'éditeur de code source XML Oxygen

### 1.2.b La problématique de l'édition de documents structurés

Au-delà des problématiques liées à la structuration de l'information, notamment au vu de la gestion de cohérence, la saisie déclarative dans un langage structuré ne se fait pas sans difficulté. En effet, l'écriture structurée en extension tend à produire des documents qui peuvent représenter un volume important. Le contenu peut alors être difficilement

appréhendable lorsqu'il est affiché sur un écran. De même, la profusion des balises peut en rendre la lecture difficile.

De là, bien que nous ayons évoqué le fait que l'analyste ait pu exploiter l'éditeur de code source, il n'en reste pas moins que la visualisation des contenus sous forme d'un code balisé a introduit des phénomènes de désorientation. Le système de représentation (code balisé) qui était appréhendable par le non-informaticien quand le contenu était réduit, est devenu problématique quand celui-ci s'est étendu. La visualisation en code source indenté ne constituait alors plus une métaphore efficace au sens où le système sémiotique ne permet plus d'appréhender toutes les composantes de l'information en cours de représentation. De ce fait, nous avons été conduits à adapter l'infrastructure logicielle de formalisation pour proposer des alternatives en terme de mode de visualisation.

### 1.3. Infrastructure logiciel de formalisation

#### *1.3.a Editeur Wysiwym XML : donner une forme aux balises*

Tout d'abord, nous avons introduit dans le dispositif logiciel un éditeur XML que l'on peut qualifier de « Wysiwym » (cf. Chapitre 8 -1.3. page 152). Il s'agit de l'éditeur Morphon (cf. Figure 58). Cet éditeur permet en effet de traduire les éléments structuraux, ainsi que le contenu associé, composant le document (les balises) en une mise en forme particulière. Ce mécanisme est basé sur la définition d'une feuille de style CSS (Cascading Style Sheets). Cette feuille de style spécifie la traduction du code source XML vers une forme de visualisation : elle définit la sémantique visuelle du document traduisant le contrat de lecture. Morphon propose une solution permettant de paramétrer cette feuille de style via une interface ne nécessitant pas une expertise technique particulière.

Cette liberté, dans la définition de la forme de représentation de l'information, a permis à l'analyste, tout en conservant une information très structurée, de produire la visualisation qu'il souhaitait : il garde une part d'autonomie dans la définition de la forme. Par exemple, il lui était possible d'adapter cette forme à son activité et de mettre en valeur les parties du document sur lesquelles il souhaitait se concentrer durant une révision du contenu. Bien évidemment, cette nécessité de mettre en place un éditeur « Wysiwym » souligne les limites d'une approche où des utilisateurs non informaticiens devraient éditer directement en code source leurs documents. Ceci met l'accent sur la problématique de la forme et le manque évident d'outils efficace et adaptés à la conception et la construction de documents structurés. On notera au passage que l'éditeur Morphon n'est aujourd'hui plus disponible du fait de

l'arrêt de son maintien. Il est alors possible de se tourner vers une solution libre, comme Amaya (Quint et Vatton 2004) dans la poursuite de cette étude ou encore la dernière version de Oxygen (v.9) qui propose à présent une mise en forme du code source via une feuille de style.

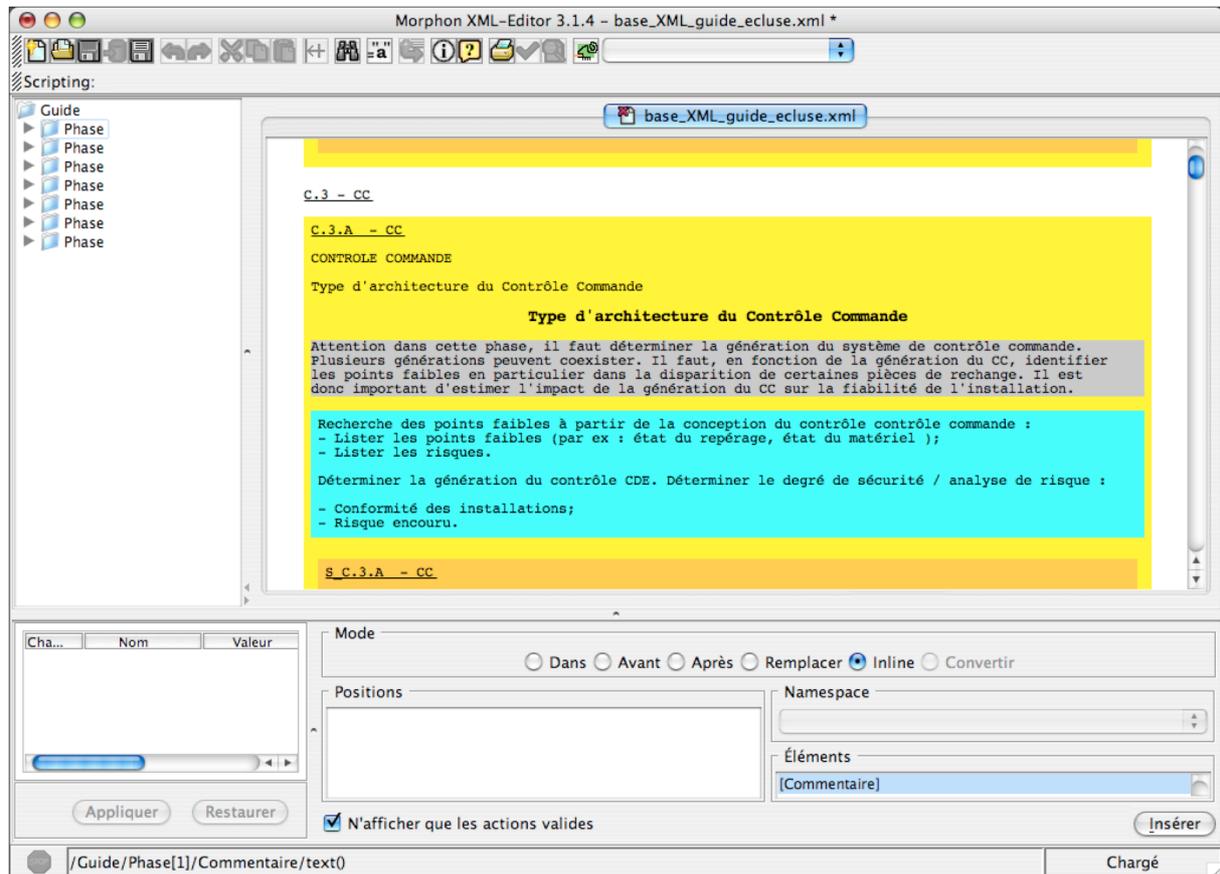


Figure 58 Interface de saisie proposée par Morphon

L'éditeur Morphon s'est révélé particulièrement adapté à traiter le contenu des briques élémentaires, mais pas nécessairement adapté à des modifications de structures importantes. Initialement, la structure a bien été construite au travers de la saisie déclarative dans l'éditeur de code source Oxygen, permettant la création d'éléments structuraux et de contenu de concert. Toutefois, nous avons évoqué la difficulté d'utiliser cet éditeur dès lors que le contenu prend de l'ampleur.

Pour répondre aux difficultés d'édition en code source, nous avons alors étendu notre dispositif de formalisation en introduisant la modalité graphique comme mode de saisie. Dans cette orientation, nous avons proposé à l'analyste deux solutions :

- Un éditeur de carte conceptuelle qui se fonde sur une visualisation arborescente d'informations.

- L'utilisation d'un tableur (Excel) qui mobilise les 2 dimensions du plan et offre une représentation matricielle.

### 1.3.b Editeur de carte conceptuelle pour une écriture graphique

Au travers de la modalité graphique, la tentative est d'offrir une représentation de l'information permettant une meilleure appréciation de l'organisation structurale et topologique de l'information. Plus généralement, la représentation graphique permet d'obtenir une vision plus globale du document. Le document structuré correspondant à une information organisée hiérarchiquement : l'arbre est la métaphore visuelle qui s'impose. Nous nous sommes alors appuyés sur l'éditeur de carte conceptuelle Freemind (<http://freemind.sourceforge.net>). Cet éditeur est fondés sur une visualisation d'information sous la forme d'une carte conceptuelle : il s'agit d'informations liées hiérarchiquement selon un nœud racine. Cette éditeur met à disposition de nombreux dispositifs qui permettent de modifier la mise en forme matérielle de l'arbre : coloriage, représentation iconique des nœuds ou des liens (cf. 0 page 292). Dans notre exploitation de Freemind, nous avons défini un algorithme générique (sous la forme d'une feuille de transformation XSL) permettant de transformer un document structuré sous la forme d'une carte conceptuelle. On se positionne à un niveau d'abstraction assez haut pour gager de la généralité : on traduit les éléments contenus dans le document structuré et leur organisation au travers des possibilités de l'architexte cible qui dans le cas présent est freemind (cf. à la section de ce chapitre 2 - page 180 pour plus d'éléments relatifs à la restitution). Cet algorithme permet également de faire la transformation inverse. Nous produisons ainsi une carte conceptuelle selon un format tout à fait standardisé qui permet de travailler aisément sur la structure du document au travers d'une représentation graphique plus adaptée à la saisie structurale de l'information. Nous imposons un principe de base dans l'utilisation de freemind. Il s'agit de définir les balises sous la forme de nœud de la carte conceptuelle, leur contenu étant construit de la manière suivante : « Nom balise | Contenu Balise ».

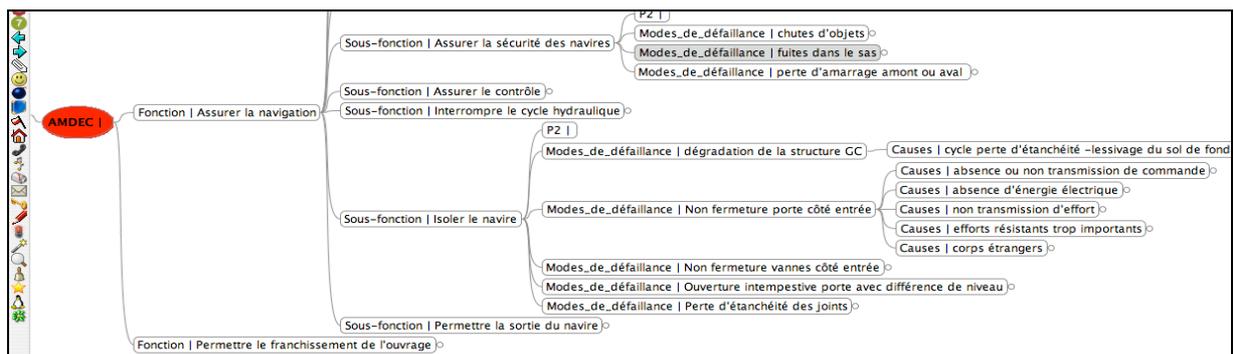


Figure 59 Carte conceptuelle Freemind calculée sur la base d'un fichier XML

### 1.3.c Tableur Excel et OpenOffice

L'usage du tableur est tellement courant dans notre contexte d'étude qu'il est impensable de vouloir écarter cette solution à la fois de saisie et de diffusion de l'information. Nous avons introduit le cas du tableur Excel et de son utilisation dans Chapitre 8 -1.2. page 151, en insistant en particulier sur le fait qu'il était d'avantage employé comme un éditeur de structuration et de mise en forme de données que comme un outil de calcul.

Le tableur Excel a donc été intégré à l'infrastructure de capitalisation. Il propose un espace matriciel de saisie dont la brique élémentaire est représentée par la cellule d'une feuille de calcul. Nous avons ainsi développés des algorithmes génériques d'analyse des feuilles de calcul saisies, permettant de transformer une feuille de calcul Excel en un document structuré bien formé. Ces algorithmes tiennent compte du contrat d'écriture récurant que nous avons pu souligner au Chapitre 8 -1.2. page 151 correspondant à une écriture hiérarchique de l'information. Ce contrat d'écriture traduit le passage d'une colonne à une autre dans la feuille de calcul comme un passage d'un niveau de profondeur à un autre dans une hiérarchie. La définition des différentes branches de la hiérarchie est identifiée par le fait que certaines cellules résultent de la fusion de cellules à différentes lignes consécutives sur une même colonne. Nos algorithmes ne vont pas plus en avant dans l'analyse des documents Excel produits, laissant de côté par exemple une mise en forme matérielle spécifique du texte ou des cellules.

En s'appuyant sur Microsoft Excel, on bénéficie de son déploiement et du fait que les utilisateurs y sont habitués. La contrepartie est que la solution mise en œuvre n'est pas, en soi, pérenne, le codage informatique des fichiers Excel étant un format propriétaire. Pour passer outre cette difficulté, nous avons initié le développement de solutions de saisie basées sur l'utilisation d'une solution libre : le OpenCalc proposé dans la suite OpenOffice (<http://fr.openoffice.org>). Le format libre que propose cette suite bureautique, qui aujourd'hui est totalement normalisé, facilite la création de documents de manière automatique ainsi que l'analyse de leur contenu lorsqu'ils sont créés de manière manuelle. Un exemple de ce type de document est présenté sur la Figure 60.

La solution de saisie via un tableur, qu'il s'agisse de Excel ou OpenOffice semble très appréciée. Elle s'intègre aisément dans la pratique de nos utilisateurs qui sont habitués à ce type de logiciels et d'interfaces. La forme matricielle de la feuille de calcul semble également une solution très adaptée à la saisie de données structurées, et de XML en particulier, si l'on

considère qu'une solution similaire est reprise dans certains éditeurs spécialisés. Ainsi, la dernière version de l'éditeur de code source XML Oxygen propose un mode de visualisation des fichiers édités sous la forme d'une grille de saisie, telle que présentée sur la Figure 61. Cette grille est construite sous une forme très similaire au contrat d'écriture identifié dans Excel tel qu'évoqué précédemment. Ceci semble confirmer l'intérêt de cette forme d'écriture des documents structurés hiérarchiques au travers d'un espace matriciel de visualisation.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
21					Fiche de visite : Identifiant : E.4.A - HM Domaine : HYDRO-MECANIQUE Nom_Court : Charpente porte levante - Sas non chômé Nom_long : Porte levante aval Charpente - Sas non chômé Commentaire : Description : Images : Check_List : Sorties :											
22					Fiche de visite : Identifiant : E.4.B - HM Domaine : HYDRO-MECANIQUE Nom_Court : Mécanisme de manoeuvre porte levante - Sas non chômé Nom_long : Porte levante aval Mécanisme de manoeuvre - Sas non chômé Commentaire : Nécessite un nettoyage des structures - Plusieurs dimensionnements en vigueur - Prévoir des manoeuvres de porte. Description : Images : Check_List : Sorties :											
23					Fiche de visite : Identifiant : E.4.C - HM Domaine : HYDRO-MECANIQUE Nom_Court : Essai Porte levante - Sas non chômé Nom_long : Porte levante aval Essai - Sas non chômé Commentaire : L'objectif de l'essai est de s'assurer que l'équipement est apte à accomplir sa fonction. Cet essai peut permettre de détecter des défaillances cachées, ou événement. Description : Images : Check_List : Sorties :											
24					Etape : Identifiant : E. Nom_Court : Z Nom_long : Z Commentaire :											
25					Fiche de visite : Identifiant : E.4.D - HM Domaine : HYDRO-MECANIQUE Nom_Court : Guidage porte levante - Sas non chômé Nom_long : Porte levante aval Guidage - Sas non chômé Commentaire : Description : Images : Check_List : Sorties :											
					Fiche de visite : Identifiant : E.4.E - HM Domaine : HYDRO-MECANIQUE Nom_Court : Blanchéité porte levante - Sas non chômé Nom_long : Porte levante aval Blanchéité - Sas non chômé Commentaire : Description : Images : Check_List :											

Figure 60 La vision métier du guide de diagnostic dans le tableur OpenOffice

Guide	Phase (3 rows)	Identifiant	Nom_Court	Nom_long	Commentaire	Etapes
		1 D	Observation du fonctionnement de l'ouvrage	Observation du fonctionnement de l'ouvrage	Le moyen principal de surveillance est l'observation visuelle et auditive. Il est par conséquent indispensable de lui donner la régularité, la rigueur et la précision qui s'imposent. Cette phase d'inspection est réalisée par l'exploitant	Etapes Identifiant Écoute et observation Nom_Court Écoute et observation Nom_long Écoute et observation Commentaire L'écoute et l'observation de l'ouvrage sans démontage dans les domaines : - Hydromécanique - Genie civil - Contrôle commande
		2 E	Inspection Sas non chômé	Inspection Sas non chômé	Lors de cette phase, il est nécessaire de rester au moins une journée sur l'ouvrage en situation de fonctionnement, avec un trafic	Etape (3 rows) Etapes (5 rows)

---

Figure 61 La saisie sous forme de grille d'un fichier XML dans la version 9 de Oxygen

#### 1.4. L'autonomie de l'analyste en question

A travers l'exemple de l'utilisation de Freemind, on voit toute la souplesse offerte par le passage dans un codage XML des informations. Il est alors possible de transformer les documents structurés afin de les adapter par exemple à divers architextes pour peu que leur format de fichier soit ouvert ou accessible via des bibliothèques logiciels spécialisées.

Cette souplesse se fait tout de même au détriment d'un acteur. L'analyste qui n'est pas nécessairement un informaticien. Il subit une perte d'autonomie pour ce qui est de la production de la forme du document. Il n'est plus totalement libre, dans ces choix de mise en forme matérielle en particulier. On lui impose quelques recommandations de saisie, que cela soit dans l'éditeur de carte conceptuelle ou dans l'utilisation d'un tableur. Nous avons évoqué que pour l'éditeur XML Morphon, que nous considérons comme Wysiwym, un analyste non informaticien peu construire de manière autonome sa visualisation.

Dans les autres solutions, il est alors dans l'obligation de faire appel à un informaticien maîtrisant les technologies XML, par exemple pour produire une forme de visualisation de l'information adaptée à son activité ou à l'activité des futurs utilisateurs (lecteurs) du document. Cette perte d'autonomie reste toutefois limitée. Tout d'abord, l'environnement de création de document structuré qui lui est proposé lui laisse encore un champ très libre dans le développement d'une structuration des données adaptée à ses besoins. C'était bien là un de nos objectifs premiers, que l'analyste ne soit pas obligé d'anticiper les structures de données correspondant à la modélisation. Au regard de la forme, même si dans un premier temps, et à ce jour, l'informaticien est encore incontournable, nous verrons par la suite qu'il est possible de diminuer la nécessité de son implication. Nous allons déjà apporter des éléments de réponses dans la section suivante qui aborde le problème de restitution des documents structurés. On se reportera également au Chapitre 12 -2.2. page 273, qui présente les perspectives de développement d'un atelier de valorisation des documents structurés permettant, au moins dans l'activité courante de l'analyste et pour des traitements peu complexes, de développer en autonomie ses propres stratégies de visualisation des documents.

## 2 - Document et restitution

Le processus de restitution doit permettre de transformer le contenu des documents structurés sous une forme visuelle et éventuellement interactive adaptée, en tenant compte de facteurs liés, par exemple, au contexte dans lequel elle doit s'insérer. Dans ce processus, il s'agit de

valoriser au mieux la séparation fond-forme véhiculée par le document numérique structuré, tout en se prémunissant de la dilution du document évoqué dans le Chapitre 3 -2.2. page 76. Cette dilution est liée au caractère réutilisable des documents numériques. La souplesse et la facilité de réutilisation du contenu des documents, que cela soit sur le plan de la forme et du fond tend à rendre les documents de plus en plus modifiables, pouvant faire oublier les fondamentaux du concept de document (cf. Chapitre 3 -1 - page 63). En effet, la tentation est alors grande de le considérer comme un « lego » manipulable oubliant toute la valeur du travail de l'auteur dans la réalisation d'un document.

Cette dimension auctoriale n'est pas absente dans le cadre de la manipulation de documents structurés, subissant des modifications de forme et de contenu et donnant naissance à de nouveaux documents. Reprenant le positionnement que l'interface homme machine constitue une métacommunication entre concepteur et utilisateur de ces interfaces (cf. Chapitre 4 -2.3. page 96), la dimension auctoriale dans la manipulation de documents numériques structurés se retrouve noyée dans les algorithmes implémentés. Il convient donc de les développer avec précaution de manière à conserver la trace des traitements effectués. Il s'agit également d'être en mesure d'en présenter les traces à l'utilisateur afin de lui permettre d'accéder à une dimension explicative pour une meilleure appropriation et compréhension des contenus diffusés.

## 2.1. Vision globale du processus de restitution

L'approche de restitution que nous proposons prend comme acquis que les informations à restituer sont sous la forme de documents numériques structurés. Ces documents peuvent être le résultat de l'approche de capitalisation que nous venons de présenter, ou procéder d'une extraction d'un système à base de connaissances dans le cas où l'on est face à une solution logicielle préexistante. Notons dès à présent, même si tout ce qui sera évoqué ici reste valable pour d'autres technologies existantes ou à venir, l'approche a été implémentée en s'appuyant sur le standard XML.

Le processus de restitution peut être dissocié en trois étages indépendants qui vont successivement traiter de :

- La manipulation du contenu,
- La mise en forme,
- L'ajout d'interaction.

Au niveau de *la manipulation du contenu*, il s'agit d'exploiter au mieux le fait que le document intègre des éléments structuraux définissant un découpage logique ou généraliste (cf. Chapitre 3 -1.3.b page 67) de son contenu. Le contenu est alors manipulable par la machine selon une stratégie à définir. Ainsi, cette manipulation pourra tout aussi bien faire référence à des mécanismes de filtrage et de réorganisation, ou encore la fusion de différents documents structurés.

L'étage de *mise en forme* a pour objet de transformer le document structuré brut, initialement sous la forme d'un code source balisé, vers un éventail de visualisations. Les différentes formes de visualisations doivent être construites en tenant compte d'éléments de contexte tels que le public visé, l'activité que doit appuyer la visualisation ou encore les conditions dans lesquelles cette activité pourra être lue. Il s'agit ici de mettre en relation la sémantique des informations du document structuré avec les possibilités de visualisations disponibles pour créer un système, un langage, de représentation.

A l'image de la mise en forme, l'étage *d'ajout d'interaction* peut associer aux éléments constituant le document structuré des comportements interactifs particuliers. Il pourra s'agir de liens hypertexte, de mécanismes de zoom ou de déplacements si l'on est face à une image, de masquage, etc...

Les étages de mise en forme et d'ajout d'interaction vont s'appuyer dans leur fonctionnement interne, d'une part sur une analyse des informations contenues dans le document structuré (repère 1 sur la Figure 62), et d'autre part sur une étude des possibilités en terme de visualisation et d'interaction de l'interface d'accès vers laquelle on tend (repère 2 sur la Figure 62). Ainsi, ces deux étages vont faire un appariement entre les possibilités des interfaces, et les données à restituer à travers une catégorisation des informations largement inspirée de la sémiologie graphique de Bertin.

Nous traiterons ici essentiellement du problème de la mise en forme des données, que cela soit pour définir une visualisation mais également pour une mise en adéquation avec une interface cible. Nous laisserons de côté l'ajout d'interaction bien que notre étude est abordé ce champ, il est encore prématuré de vouloir décrire cette voie d'étude.

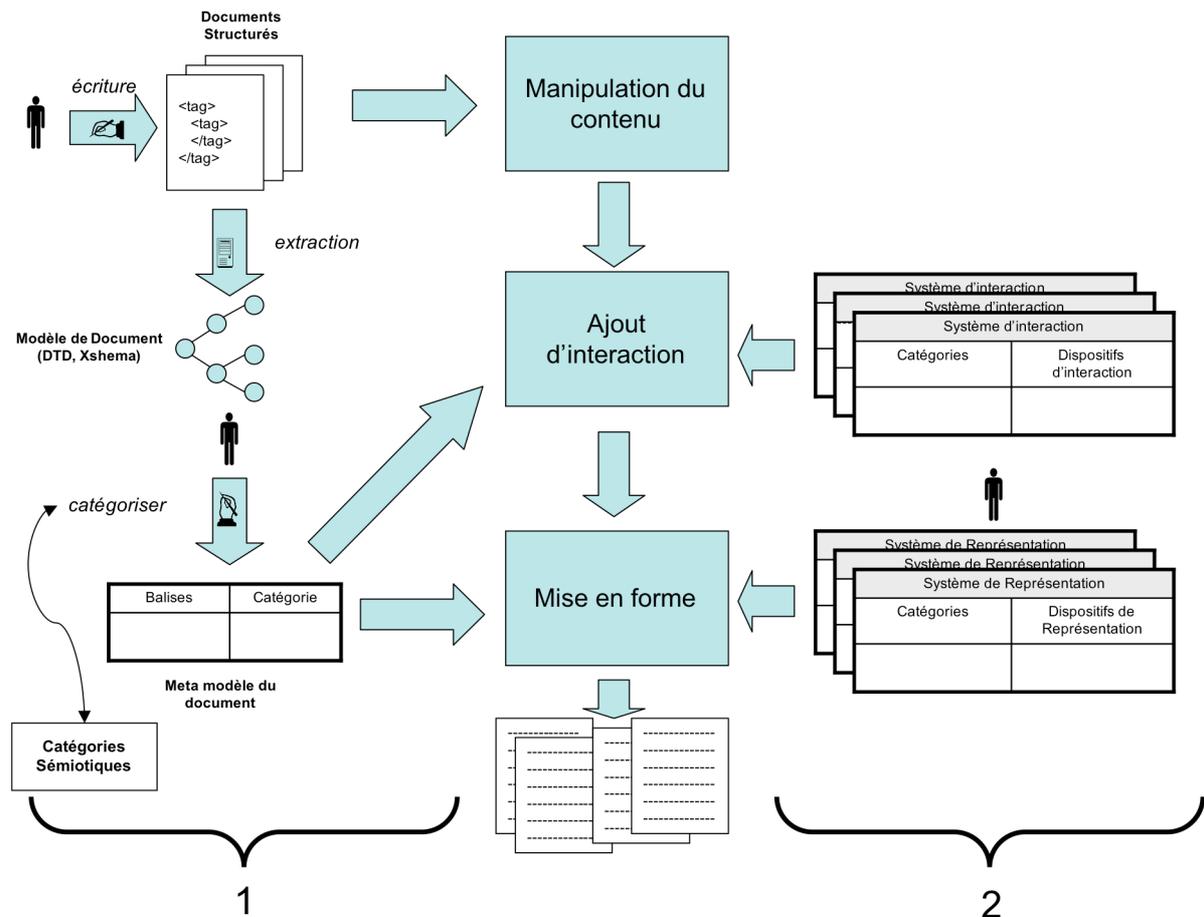


Figure 62 Processus de restitution

## 2.2. Méthode de construction d'une restitution documentaire

L'approche de restitution que nous proposons est très largement inspirée de la méthode de construction de l'image proposée par Bertin (cf. Chapitre 5 -2 - page 116). Cette méthode, liée à la sémiologie graphique, peut être synthétisée sous la forme suivante :

- Inventaire et analyse des données à représenter dans l'image,
- Pour chaque information que l'on souhaite diffuser, communiquer ou présenter pour être analysée : identification des composantes à faire apparaître,
- Pour chaque composante : caractériser le niveau de perception à laquelle elle correspond,
- Associer aux composantes une variable visuelle en fonction du niveau de perception (cf. tableau présenté page 121).

Notre proposition est fondée sur une adaptation de la méthode de construction de l'image à notre contexte d'étude et à l'objet que nous souhaitons produire. Il s'agit donc de passer d'une

méthode de construction de l'image à une méthode de construction du document qui constitue notre support véhiculant l'inscription à restituer.

Le passage de l'image au document pose plusieurs difficultés. Que le document soit édité manuellement ou calculé par un ordinateur, les architectes, interfaces et formats d'édition et de diffusion ne permettent pas toujours d'accéder aux variables visuelles pures évoquées par Bertin. En effet, ces variables peuvent tout d'abord être limitées au niveau de leur longueur. Un exemple serait de ne pouvoir utiliser qu'un nombre limité de formes dans la réalisation d'un dessin au sein d'un document. C'est notamment le cas lorsque l'on utilise un logiciel proposant une bibliothèque de formes prédéterminées.

Une deuxième difficulté est que l'image construite par Bertin se caractérise par la définition d'un langage monosémique, au sens où il s'agit d'un espace cohérent de représentation graphique dans lequel tous les éléments inscrits prennent un sens bien défini. Le document, pour sa part, correspond à un support dans lequel il va être possible d'agréger, d'associer, différents espaces de représentation. Le document est multimodal. On peut alors le considérer comme une combinaison de différentes images. Chaque image constituera à la fois un espace de représentation indépendant intégrant son langage propre. On peut illustrer ce point avec l'exemple d'un document tout à fait classique, qui intègre un contenu textuel, précédé d'un sommaire, référençant les titres contenus dans ce document. Le sommaire et la zone de contenu vont chacune représenter un espace de représentation particulier, soumis à ses règles propres. Il constitue un support d'accès à l'information. Le sommaire peut être vu comme une vision synthétique et globale du document. Il fait apparaître le découpage logique du document en faisant transiter le niveau de profondeur de chaque élément cité par une indentation ou une mise en forme spécifique. Il s'agit donc bien d'une zone dans laquelle l'écriture est cadrée par une sémantique particulière : cette dernière correspond au contrat de lecture et d'écriture. Il en est de même pour la zone de contenu où les titres pourront prendre une mise en forme matérielle particulière en fonction de leur niveau dans la hiérarchie ou encore impliquer des sauts de page. Au sein du contenu, on peut encore définir ou identifier des sous espaces de représentation particuliers tels que des diagrammes, des tableaux...

Au-delà de la sémiologie graphique et de la construction d'images fixes, nous souhaitons pouvoir valoriser l'interface numérique et intéresser à l'interactivité. Il s'agira donc dans les documents d'être en mesure d'intégrer des moyens interactifs lorsque cela est possible mais en intégrant la philosophie de construction de l'image de Bertin. L'idée est alors d'associer

certaines comportements interactifs en tentant de prendre en compte une certaine efficacité au regard du type d'information sur laquelle ils sont greffés. Nous verrons que nous n'avons que commencé à investiguer cet aspect.

L'approche que nous proposons peut être dissociée en quatre étapes :

- Définition d'un modèle de domaine
- Définition d'un modèle de perception
- Définition d'un modèle d'architexte
- Construction du document

La *définition d'un modèle de domaine* correspond à une analyse des données devant être restituées. Il s'agit de bien les identifier et éventuellement de les compléter si nécessaire afin de limiter le tacite.

La *définition d'un modèle de perception* doit permettre de caractériser les informations à visualiser au regard de la manière dont elles devront être perçues visuellement. Il s'agit de se dégager de la sémantique métier des informations dans leur contexte initial d'où elles sont issues. Par exemple, un domaine de compétence deviendra une famille d'éléments ou association d'éléments.

La *définition d'un modèle d'architexte* correspond à une analyse des possibilités de l'interface de capitalisation ou de restitution. L'objectif est d'inventorier les capacités de visualisation et d'interaction disponible.

Enfin, la *construction du document* propose une mise en relation du modèle de perception avec le modèle d'architexte. C'est à ce moment que l'on va traduire la sémantique des informations au travers des possibilités visuelles et interactives de l'interface cible, en tentant d'introduire la notion d'efficacité (cf. Chapitre 5 -2.2.a page 120).

Dans les sections suivantes, nous passons en revue les différentes étapes proposées dans notre approche, en illustrant chacune par un exemple permettant d'illustrer l'ensemble du processus de restitution documentaire.

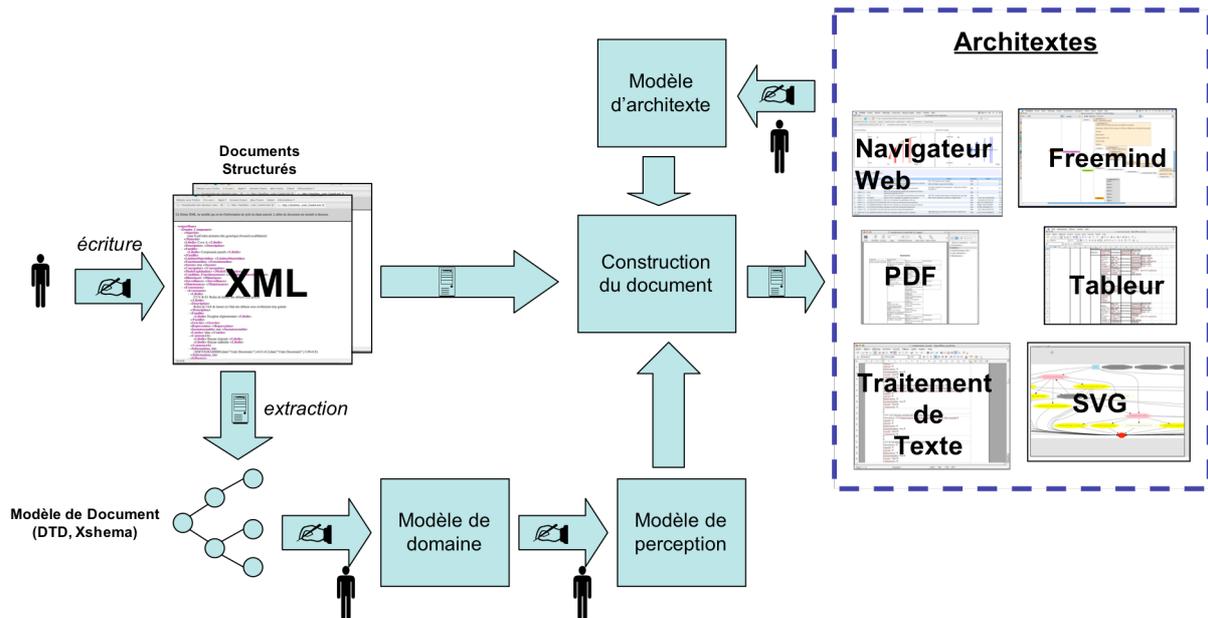


Figure 63 Approche de restitution documentaire

## 2.3. Définition d'un modèle de domaine

### 2.3.a Définition

La définition du modèle de domaine a pour but de fournir une modélisation des informations à représenter. Sa définition passe par l'*identification et l'analyse du jeu de données* que l'on souhaite visualiser. L'objet est ici d'identifier l'étendue de l'espace de données à visualiser. Il s'agit d'effectuer :

- Un inventaire de données,
- L'extraction d'informations implicites,
- Une prise en compte du contexte.

L'*inventaire des données à restituer* peut se limiter à inventorier les éléments constituant le modèle qui régit le jeu de données à restituer. Dans notre étude, il s'agit de modèle de documents structurés qui va tout à la fois préciser les éléments structuraux utilisables ( le vocabulaire) et la manière dont on peut les imbriquer, les organiser (syntaxe) (cf. Chapitre 3 - 1.3.c page 69). A ce stade, il est également possible d'élaguer certains éléments du modèle d'origine que l'on ne souhaite pas restituer.

L'*extraction d'informations implicites* doit permettre d'identifier des éléments implicites pouvant se révéler pertinents dans le processus de restitution de l'information. A ce niveau, il

s'agit de spécifier d'éventuels éléments tacites, non pas pour l'humain, mais bien pour une exploitation des données par l'ordinateur. Par exemple, dans un modèle de connaissances qui lie des « symptômes » et des « défauts » sur un matériel, les relations introduisent par exemple une notion d'orientation ou de précédence entre les éléments. Cette interprétation faite par une personne quelconque, souvent rendue visible par la représentation graphique d'un modèle, doit être transformée en un attribut supplémentaire dans le modèle de domaine afin de pouvoir être exploitée par l'ordinateur et rendre compte d'une visualisation.

La *prise en compte du contexte* est également un élément important dans la définition des modalités de restitution. Ce contexte pourra aussi bien être lié au domaine d'expertise de l'utilisateur qu'à la tâche que va instrumenter la restitution des informations. Sur la base de l'exemple précédent, selon que l'on se situe dans une activité de diagnostic – identification des causes d'une défaillance - ou de pronostic – prédiction d'une défaillance - , il pourra s'agir respectivement de faire apparaître les défauts suivis des symptômes ou inversement. Il est alors important de définir une relation d'ordre entre les éléments pouvant prendre la forme d'une indication de précédence entre eux. Le modèle de domaine rend explicite l'ensemble des informations qui vont pouvoir jouer un rôle dans la définition d'une forme visuelle et interactive

### *2.3.b Exemple : les guides de diagnostic*

Un exemple simplifié de modèle de document des guides de diagnostic est présenté sur la Figure 65. A partir de ce modèle et d'une analyse du document structuré correspondant, l'ensemble des éléments structuraux vont pouvoir faire l'objet d'une visualisation. Les éléments structuraux, correspondant aux balises, pourront alors être dissociés selon qu'ils contiennent un contenu textuel ou non, et selon qu'ils contiennent des éléments structuraux ou non. En fonction de ces critères, l'élément structurant n'est pas ou ne traduit pas un contenu de même nature :

- Les éléments structuraux « Guide », « Phase », « Etapes », « Etape » et « Fiche de Visite » constituent l'organisation logique du document, et traduisent le découpage logique du diagnostic tel qu'il a pu être capitalisé. C'est aussi bien le nom de l'élément qui est important que la zone qu'il délimite. Le nom de l'élément traduit ainsi la nature de la zone délimitée

- Les éléments structuraux « Identifiant », « Nom long », « Commentaire », « Description » et « Domaine » correspondent à un balisage considéré comme plus généraliste des informations. Il ne constitue pas des éléments du diagnostic mais tendent à caractériser la nature de leur contenu textuel.

```

- <Guide>
- <Phase Identifiant="C" Nom_long="Conception et spécificité de l'ouvrage" Nom_Court="Conception et spécificité de l'ouvrage">
- <Commentaire>
  Le retour à la conception peut s'avérer indispensable. En effet pour comprendre un état de défaillance, il est parfois nécessaire de remonter à la conception du composant mais aussi à sa technologie, des matériaux utilisés ou du dimensionnement initial. Toute modification sur l'ouvrage nécessite un retour à la conception. Conception : - Fonctions qui doivent être remplies par l'objet et conçues pour son mode d'utilisation. - Caractéristiques mécaniques qui ont permis de répondre aux fonctions. Réalisation et réception (éventuellement modification) : - Fonctions qui doivent être remplies par l'objet et conçues pour son mode d'utilisation. Modifications des fonctions conçues : - Réception en terme de composant, approvisionnement, assemblage, installation (PV de réception). Exploitation et utilisation (usure et vieillissement) : - Modes d'utilisation effectifs pouvant avoir été modifiés volontairement ou éventuellement dérive des réglages. - Mécanisme de vieillissement, incident et accidents.
</Commentaire>
- <Etapes Identifiant="Conception" Nom_long="Etude de la conception de l'écluse " Nom_Court="Etude de la conception de l'écluse ">
  <Commentaire>
- <Etape Identifiant="C.1-HM" Nom_long="Conception et spécificité de l'ouvrage HM">
- <Fiche_de_visite Identifiant="C.1.A-HM" Domaine="HYDRO-MECANIQUE" Nom_long="Analyse de la conception de l'ouvrage hydro-méca" Nom_Court="Analyse de la conception de l'ouvrage hydro-méca">
  HYDRO-MECANIQUE
  <Commentaire>
    Le dossier de construction est un dossier comprenant notamment : - Le dossier administratif ( concession et avenant, cahier des charges particulières concernant l'ouvrage), - Les plans conformes à exécution, - Le rapport de fin de chantier, - Les rapports géologiques (reconnaisances, réception de fond de fouille, suivi d'exécution de travaux), - Les notes de calcul (pour permettre de connaître les hypothèses), - Les descriptifs techniques, - Le dossier photographique, - Le rapport de mise en eau et en particulier, les incidents éventuels survenus lors de celle-ci ainsi que les moyens mis en oeuvre pour y remédier. Les documents de la vie de l'ouvrage, ces documents constituent un dossier qui contient principalement les éléments suivants : - Les rapports d'études et de travaux d'entretien, de modification et de comportement de l'ouvrage (plans, descriptifs techniques et quantitatifs, faits significatifs des études et du chantier), - Les fiches de visites et les rapports d'expertise éventuels, - Le cas échéant, les rapports d'essais et de mesures effectués {hydrauliques, caractéristiques des matériaux, sondages...). - Le cahier des charges de la concession ou un extrait de la partie relative au canal. - Le règlement général de la police de navigation. - Le code des voies navigables et de la navigation intérieure. - Le code de navigation propre à la voie d'eau considérée. - La liste des jours légaux de chômage de la navigation. - Fiche ouvrage et dossier ouvrage
  </Commentaire>
- <Description>
- <Inspection>
  Recherche des points faibles à partir de la conception du sas ou des organes : - Contrôler le poids de la vanne et refaire les calculs de dimensionnement - Conception des tiges des vannes d'aqueduc; - Conception du mécanisme de manoeuvre de la porte aval et amont; - Câble d'équilibrage: - Chaînes. Dimensionnement de l'écluse et trafic fluvial : - Evolution du dimensionnement des bateaux: - Défaut de

```

Figure 64 Le guide sous la forme d'un document structuré en XML

Dans le cas qui nous occupe ici, on peut identifier une information implicite dans le document structuré. Le diagnostic introduit une séquentialité dans les opérations qui permettent de le mener à bien. Dans la rédaction du document structuré cette séquentialité est implicite puisqu'elle correspond à l'ordre dans lequel les éléments de diagnostic ont été écrits dans le document. La sémantique associée à cet ordre reste implicite au premier abord et il s'agira de la prendre en compte ou d'ajouter une information permettant par exemple d'identifier la précedence entre les éléments de diagnostic.

Une autre information implicite, inhérente à tout document structuré, correspond à l'imbrication des balises qui traduit l'arborescence du document structuré traité. Cette imbrication pour le cas du guide ne traduit pas la même information selon les balises. Ainsi, pour les éléments de diagnostic de type Phase, Etapes, Etape et Fiche de visite, elle traduit le découpage du diagnostic et sa logique d'enchaînement. Pour les autres éléments représentés comme des attributs sur le modèle, l'imbrication traduit leur relation de caractérisation à l'élément de niveau supérieur. Ainsi, le contenu d'un segment délimité par une balise « Identifiant » caractérise de manière unique les « Phase », « Etapes », « Etape » et « Fiche de

visite ». On aboutit ainsi à partir du modèle de document au modèle « entité-relation » présenté sur la Figure 64.

Nous pourrions également prendre en compte d'autres informations telles que la profondeur des nœuds dans le document structuré ou encore l'introduction de liens avec d'autres documents.

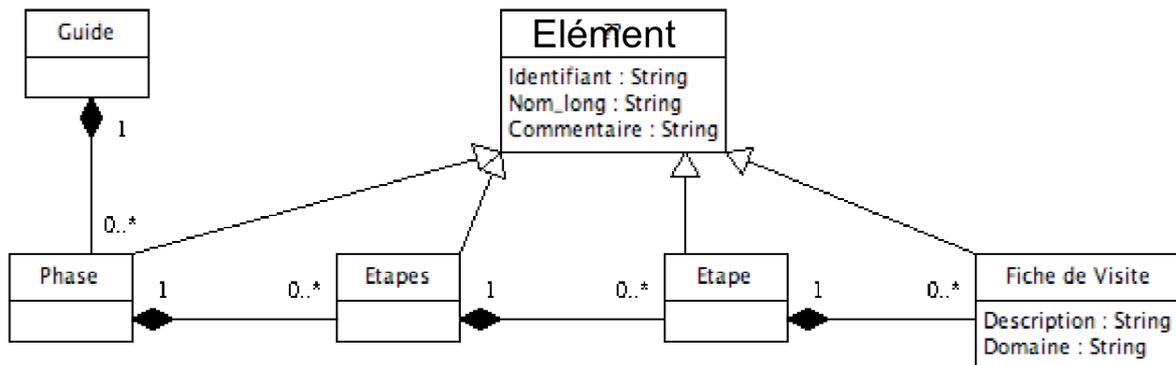


Figure 65 Modèle conceptuel du guide de diagnostic

## 2.4. Définition d'un modèle de perception

En sémiologie graphique, une *information* est définie comme « une série de correspondances entre un ensemble fini de concepts de variation ou de composantes » (Bertin 1973)(cf. Chapitre 5 -2.1. page 117). Plus simplement, une information va être le résultat d'une image. Bertin évoque l'information comme un contenu traductible d'une pensée dans la représentation graphique. Par exemple, sur un graphique à deux axes, l'information se situe au niveau des points qui traduisent la correspondance entre deux valeurs (appartenant à un concept de variation ou une composante) portées sur les axes.

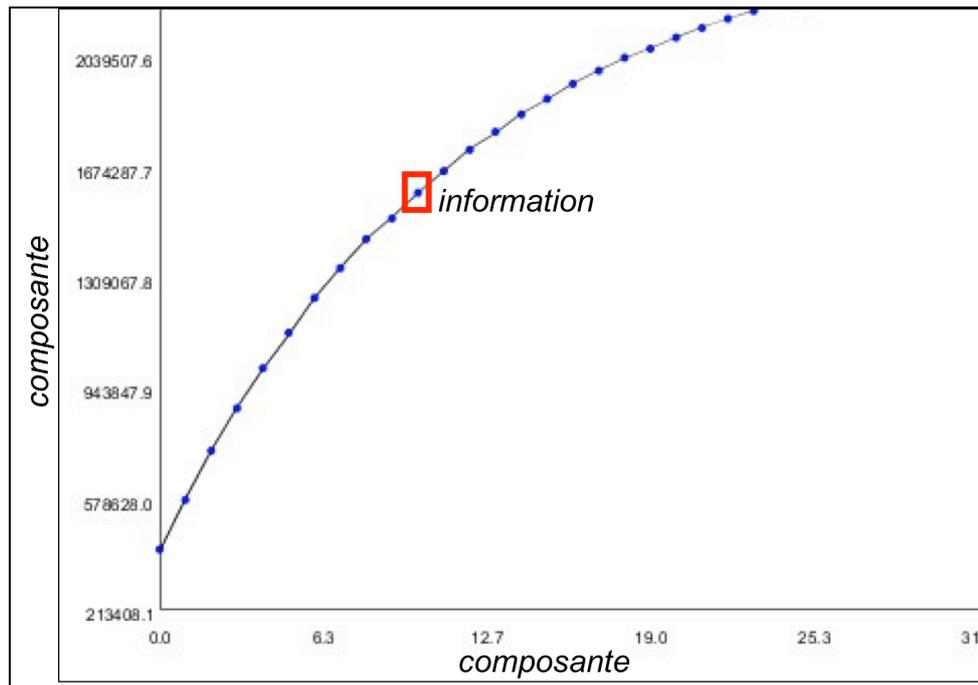


Figure 66 Un exemple de graphique à 2 dimensions

Notons au passage que les informations produites à la croisées de plusieurs composantes vont pouvoir elles-mêmes représenter des éléments appartenant à une composante en fonction du niveau de lecture de l'image où l'on se situe. Dans une représentation sous la forme de graphe, telle que présentée sur la Figure 67, chaque nœud constitue une information qui correspond au croisement de trois composantes traduites par trois variables visuelles : le texte, la forme du nœud, la couleur du nœud. Ces nœuds qui sont des informations, à un niveau plus global s'insèrent dans une représentation de type réseau (cf. Chapitre 5 -2.2.b page 122) et appartiennent alors à une composante. Dans ce réseau, la construction graphique par les liens permet de visualiser de nouvelles informations.

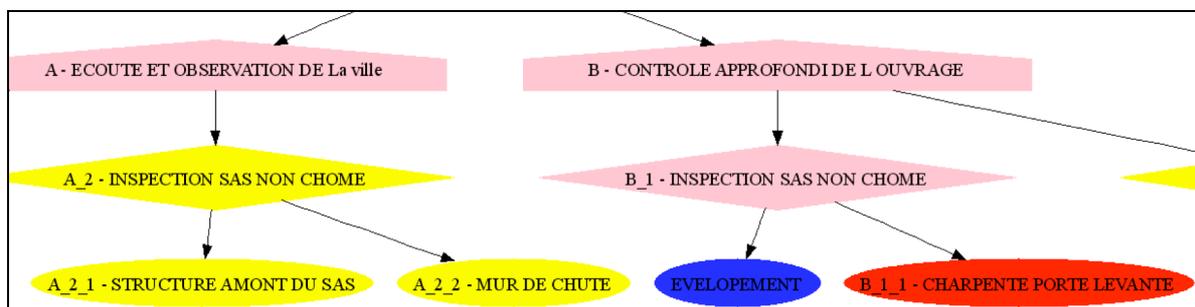


Figure 67 Exemple de visualisation sous forme de graphe

#### 2.4.a *Le modèle de perception*

Dans notre approche, il s'agit donc, dans la construction d'une représentation, de définir si les éléments structuraux délimitent des informations ou des composantes. Bien évidemment, cette caractérisation n'est valable que pour un niveau de la représentation, telle qu'on a pu le présenter à l'instant avec l'exemple du graphe. Il s'agirait donc, dans une implémentation complète de la méthode, de travailler de manière itérative et par niveau dans le document structuré.

Il s'agit ensuite de caractériser les composantes dans ce que nous nommons des catégories sémiotiques, au sens où il s'agit d'identifier la manière dont elles peuvent être perçues. Nous en définissons cinq en nous inspirant des niveaux de perception visuels proposés par Bertin (cf. Chapitre 5 -2.2.a page 120) :

- Identifiant
- Description
- Qualitative
- Ordonnée
- Quantitative

La catégorie *Identifiant* va caractériser les entités élémentaires d'un document structuré correspondant à une composante dont les éléments sont uniques. Un élément inclu dans cette catégorie doit pouvoir à la fois être perçu dans la visualisation comme unique. Son caractère unique permet de l'exploiter pour localiser une donnée permettant d'appuyer un mécanisme de lien hypertexte par exemple : l'identifiant deviendra alors la cible du lien.

La catégorie *Description* identifie les entités élémentaires constituées d'un contenu, le plus souvent textuel, mais pouvant également correspondre à des images ou des photos. Ces contenus sont difficilement caractérisables dans une catégorie perceptive car ils initient une situation sémiotique complexe lors de leur lecture.

Nous proposons ensuite la catégorie *Qualitative*, qui renvoie aux entités élémentaires du document, définissant la qualité, la nature d'une information. Cette catégorie qualitative est en relation directe avec les perceptions associatives et sélectives qui permettent le regroupement ou la dissociation d'ensembles d'éléments dans une représentation (cf. Chapitre 5 -2.2.a page 120).

La catégorie *Ordonnée* renvoie aux entités élémentaires qui délimitent un contenu ordonnable au sein de la composante à laquelle ils appartiennent. En reprenant la notion d'ordre évoquée en sémiologie graphique, les éléments doivent pouvoir être triés naturellement, de manière intuitive, sans nécessiter de devoir définir une règle particulière. C'est par exemple le cas des lettres de l'alphabet, qui s'ordonnent naturellement, à condition d'être issue d'une société, une culture, exploitant cet alphabet.

Enfin, la catégorie *Quantitative* correspond à des éléments structuraux du document qui délimitent des contenus relevant de valeurs numériques. On doit pouvoir effectuer le rapport entre deux valeurs permettant d'identifier une différence, une distance et ce sans faire appel à une légende.

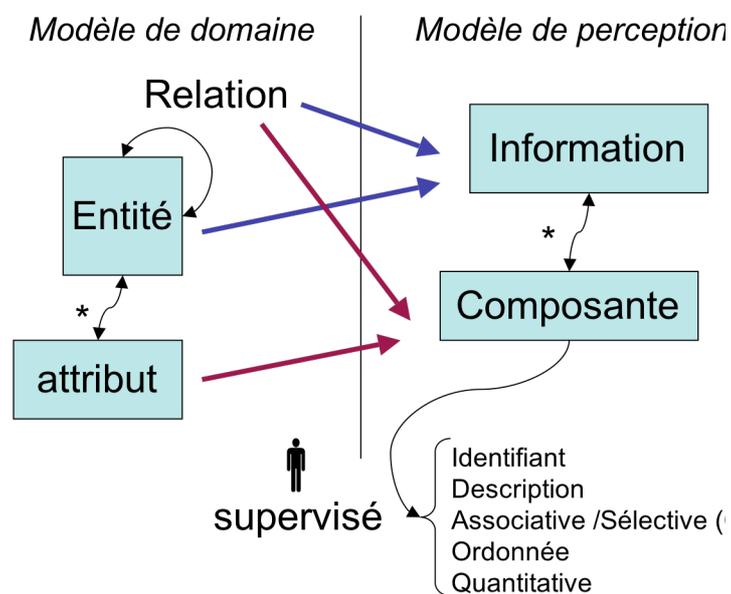


Figure 68 Du modèle de domaine au modèle de perception

#### 2.4.b Du modèle de domaine au modèle de perception : les guides de diagnostic

Dans une version simplifiée des guides de diagnostic, on peut identifier les éléments structuraux suivants : Guide, Phase, Etapes, Etape, Fiche de visite, Identifiant, Nom long, Commentaire et Description.

Nous ajoutons, dans la définition du modèle de domaine, le nom des éléments structuraux (« Nom de balise ») qui constitue une information importante au regard de notre méthode de capitalisation. En effet, les éléments structuraux vont traduire un balisage logique ou généraliste. Le balisage logique a tendance à définir des entités conceptuelles nécessaire à la

modélisation, alors que le balisage généraliste tend à identifier et caractériser sémantiquement le contenu délimité.

Nous proposons dans le tableau 3 une catégorisation des éléments constituant le modèle de domaine pour passer au modèle de perception. Ainsi, les éléments structuraux « Guide », « Phase », « Etapes », « Etape » et « Fiche de visite » sont catégorisés comme des Informations. Il s'agit en fait d'éléments structuraux traduisant un balisage logique qui définit un découpage du diagnostic en plusieurs niveaux. Les autres balises sont des composantes.

Les éléments « Identifiant » et « Nom long » sont considérés comme des composantes des Informations que nous venons de spécifier permettant de les identifier de manière unique.

Les éléments « Commentaire » et « Description », qui correspondent à des explications textuelles dépassant l'intitulé factuel, sont catégorisés comme des Descriptions.

Enfin, l'élément « Domaine » ainsi que les noms de balises sont catégorisés comme des composantes qualitatives des informations précitées. La balise « Domaine » permet de préciser le domaine de compétence auquel renvoie la partie de diagnostic dans laquelle on se situe. Le nom de la balise constitue une composante qualitative au sens où elle spécifie des intitulés dans le découpage du diagnostic. « Domaine » et nom de balise constitue ainsi des composantes permettant d'effectuer des regroupements, créant ainsi des familles d'éléments.

Tableau 3 Catégorisation des éléments structuraux et données indentifiables dans l'exemple des guides de diagnostic

<b>Éléments structuraux du document</b>	<b>Catégorie</b>
Guide, Phase, Etapes, Etape, Fiche de visite	Information
Identifiant, Nom long	Identifiant
Commentaire, Description	Description
Domaine, Nom balise	Qualitative

## 2.5. Modèle d'architexte

La sémiologie graphique de Bertin (Bertin 1973) que nous avons présentée dans la revue de littérature ( cf. Chapitre 5 -2 - ) illustre une méthode de construction de l'image permettant de mettre en relation des informations avec les moyens de visualisation du système graphique pour aboutir à une représentation appuyant le traitement de l'information. La sémiologie graphique reste une référence dans la définition d'image mais il peut être difficile de totalement l'appliquer dans un contexte réel.

En effet, nous avons évoqué le fait que nous ne souhaitons pas développer une plate-forme « tout en un » permettant de couvrir l'ensemble de la démarche de gestion des connaissances de la capitalisation à la restitution. Notre proposition, que nous pensons plus flexible, et surtout plus pérenne, est de nous appuyer sur des logiciels et interfaces existants qui sont, ou dont on peut penser qu'ils peuvent être, déployés sans trop de difficultés.

Cette approche limite par contre les marges de manœuvre au regard de la définition de la forme, puisqu'il s'agit de s'adapter aux possibilités offertes par l'interface de lecture ou d'écriture choisie, ce qui va contraindre le résultat final. En effet, les interfaces existantes ne mettront pas toujours à disposition des variables visuelles aussi pures que celles évoquées par Bertin, ou alors elles pourront imposer quelques restrictions.

#### 2.5.a Le modèle d'architexte

Nous mettons, sous l'intitulé architexte, les logiciels ou interfaces sur lesquels nous souhaitons nous appuyer. Nous l'avons évoqué au Chapitre 3 -3.2.a page 78, un architexte est défini comme un « *logiciel considéré dans sa capacité à mettre en forme et conditionner l'écriture. L'architexte (traitement de texte, navigateur, logiciel de présentation, etc.) procède à une écriture de l'écriture : il relève de la production de formes qui se situent en amont de l'acte d'écriture-lecture et en déterminent les conditions de possibilité (du grec, origine et pouvoir)* » (Jeanneret et Tardy 2007). Ainsi, dans notre étude, l'architexte peut tout aussi bien correspondre à un logiciel ou une interface auteur ou de diffusion de l'information, qu'à un format ou une bibliothèque spécifique de visualisation d'information. Il se caractérise par les « capacités de mise en forme » et « le conditionnement de l'écriture » qui déterminent les « conditions de possibilités » de lecture et écriture

L'architexte se définit tout d'abord par les espaces de représentation qu'il propose. Un espace de représentation va correspondre à un mode particulier de visualisation ou d'interaction avec l'information. Pour chaque espace de représentation, on pourra ainsi identifier à la fois des moyens de visualisation et des moyens d'interaction disponibles et utilisables : ceci correspond aux capacités de mise en forme. Leur utilisation n'est pas toujours entièrement libre puisqu'elle peut être conditionnée par l'implantation de l'architexte.

Ainsi, chaque espace de représentation peut véhiculer ses propres règles au niveau de la représentation de l'information qui vont conditionner écriture et visualisation. Ces règles peuvent être prédéfinies lorsque l'on s'appuie sur un logiciel, dont le fonctionnement est peu

modifiable : c'est le cas si l'on s'appuie sur un traitement de texte. Elles peuvent être définie entièrement à la conception de la restitution si l'architexte employé offre une grande marge de manœuvre : les bibliothèques logicielles permettent de se placer dans cette situation au prix d'un effort de programmation.

En outre, ces règles, qui peuvent conditionner l'utilisation des moyens de visualisation et d'interaction, peuvent transparaître au travers de dispositif. Le dispositif correspond alors à une configuration particulière de moyens.

Avant de s'intéresser aux moyens de visualisation et d'interaction proposés par chaque architexte, il s'agit d'identifier les différents espaces de représentation qu'ils proposent. L'architexte pourra en être composé de plusieurs chacun introduisant des possibilités différentes de restitution.

Il s'agira également de dégager les contraintes liées, inhérentes à chaque espace. Ces contraintes peuvent être liées à :

- L'utilisation des moyens de visualisation et d'interaction
- Des combinaisons de moyens de visualisation et d'interaction que nous nommerons dispositifs
- Un contrat d'écriture particulier au travers des moyens disponibles

Au final, à l'image de la Sémiologie graphique, nous proposons de mettre en relation les dispositifs identifiés ou définis, et en particulier les variables visuelles ou interactives qu'ils véhiculent, avec les catégories sémiotiques évoquées dans la section précédente. Ceci doit permettre de créer un modèle d'architexte qui caractérise le fait que tel ou tel dispositif, ou tel ou tel moyen utilisable au niveau de l'architexte, est adapté pour faire transiter tel type d'information. Il s'agit de bien analyser les architextes ou interfaces que nous ciblons. Cette analyse doit à la fois permettre d'identifier la marge de manœuvre inhérente à chaque architexte, tout en identifiant bien les contraintes et les limites.

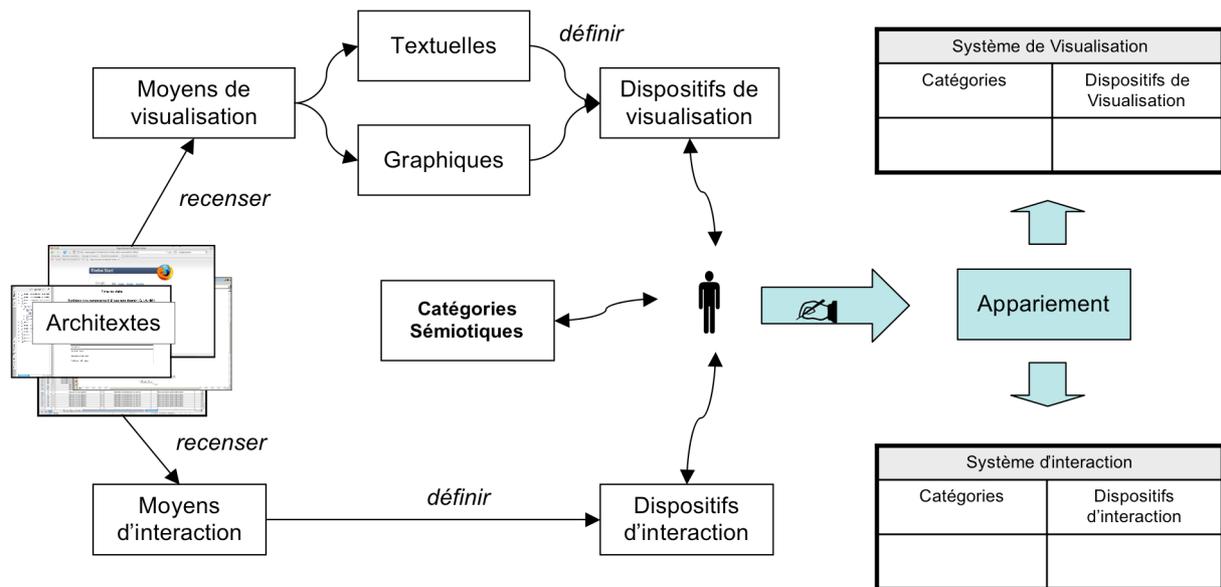


Figure 69 De l'analyse de l'architexte à la définition de dispositifs

Dans les sections suivantes, nous illustrons les différentes notions évoquées ici au travers d'exemples.

### 2.5.b Analyse des architextes

Nous nous limitons ici à une analyse simplifiée relative au cas :

- D'un lecteur de fichier de pdf
- La génération de graphe en SVG via la bibliothèque Graphviz

Une analyse plus complète de différents architextes ou formats cibles est accessible dans l'0 page 292.

#### Le cas du PDF

Dans le format pdf et des lecteurs associés, on peut identifier deux espaces principaux de représentation d'information :

- La zone de signets
- La zone de visualisation du document

La *zone de signets* est un espace de représentation qui se traduit par une visualisation arborescente d'informations textuelles. Chaque élément de l'arborescence correspond à un dispositif figé que nous nommerons « signet ». Il s'agit d'un espace de représentation très contraint, puisque les moyens de visualisation et d'interaction sont imposés en dehors de l'intitulé textuel des signets et de la cible du lien hypertexte intégré à chaque signet. Il s'agit

donc d'un espace de représentation dont le contrat d'écriture est précoordonné par la hiérarchie. La mise en forme des signets n'est pas personnalisable. Chaque signet intègre un lien hypertexte unique et permet d'emboîter les signets qu'il contient.

Le *document* constitue pour sa part un espace de représentation beaucoup plus ouvert. En effet, il peut contenir texte et image. L'ensemble des variables visuelles proposées par Bertin peuvent être exploitées. On peut ainsi exploiter les dimensions du plan dans les limites de taille du support défini, ainsi que varier taille (taille de police), forme (type de police d'écriture) ou encore couleur du texte. Les moyens d'interaction exploitables sont les liens hypertextes, qu'il est possible de placer sur n'importe quel élément contenu dans le document, et éventuellement l'annotation. Le contrat d'écriture et de lecture dans le document est totalement défini par l'auteur. Il s'agit d'un espace peu contraint dans lequel on peut définir les dispositifs comme on le souhaite.

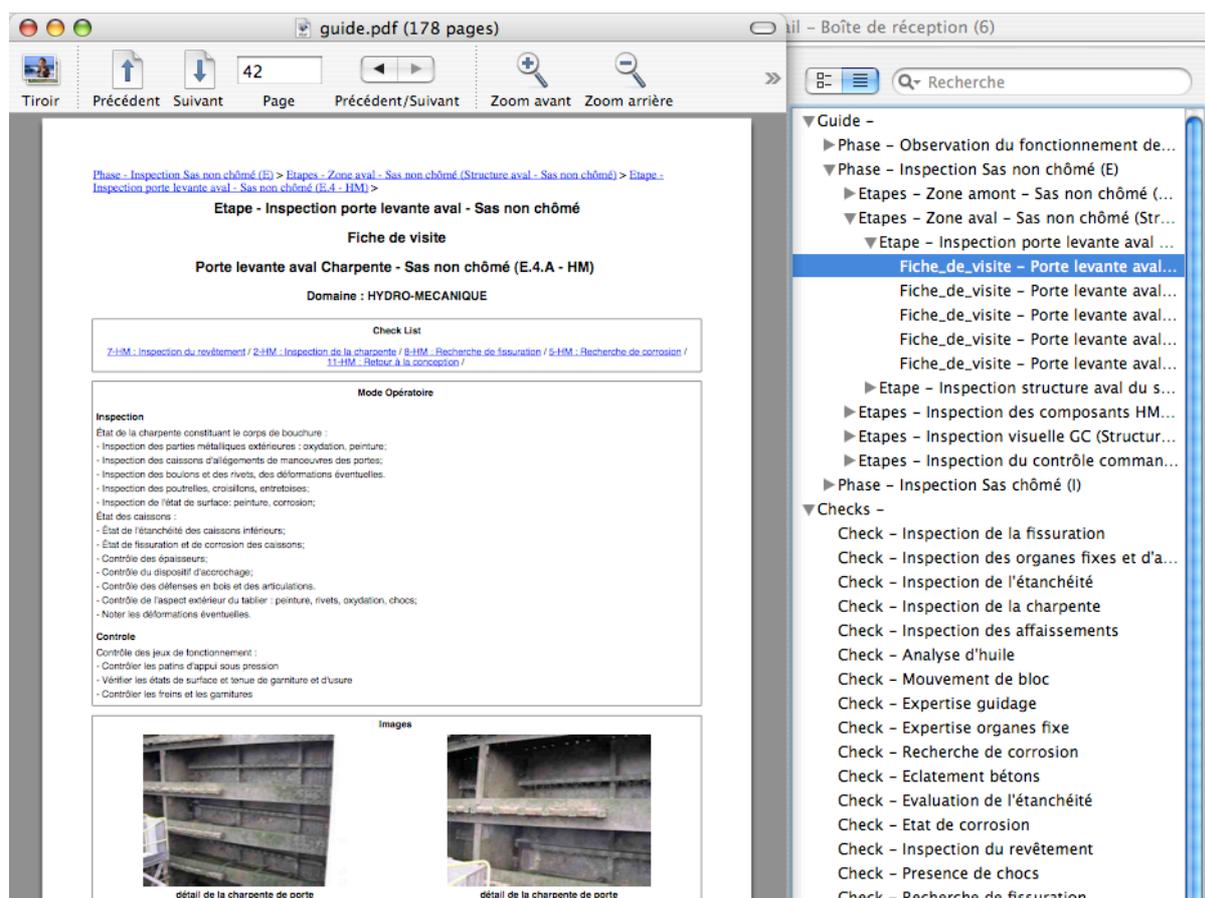


Figure 70 A gauche la zone de visualisation de documents et à droite la zone de signets

*Le cas du SVG*

La génération d'images dans le format SVG (Scalable Vector Graphic) n'a en théorie de limite que l'imagination et la créativité. Le format SVG reste toutefois difficile à manier, en dehors de l'utilisation d'un éditeur dédié permettant une construction manuelle de l'image (ex : via un éditeur SVG tel que Inkscape ou Adobe Illustrator), ou d'une bibliothèque de génération. La visualisation de graphes est un problème très étudié qui a donné lieu au développement de nombreuses bibliothèques logicielles de génération. Graphviz en est un exemple qui permet la génération de graphes dans un espace de représentation avec un plan à 2 dimensions dont la longueur est illimité (cf. Figure 71).

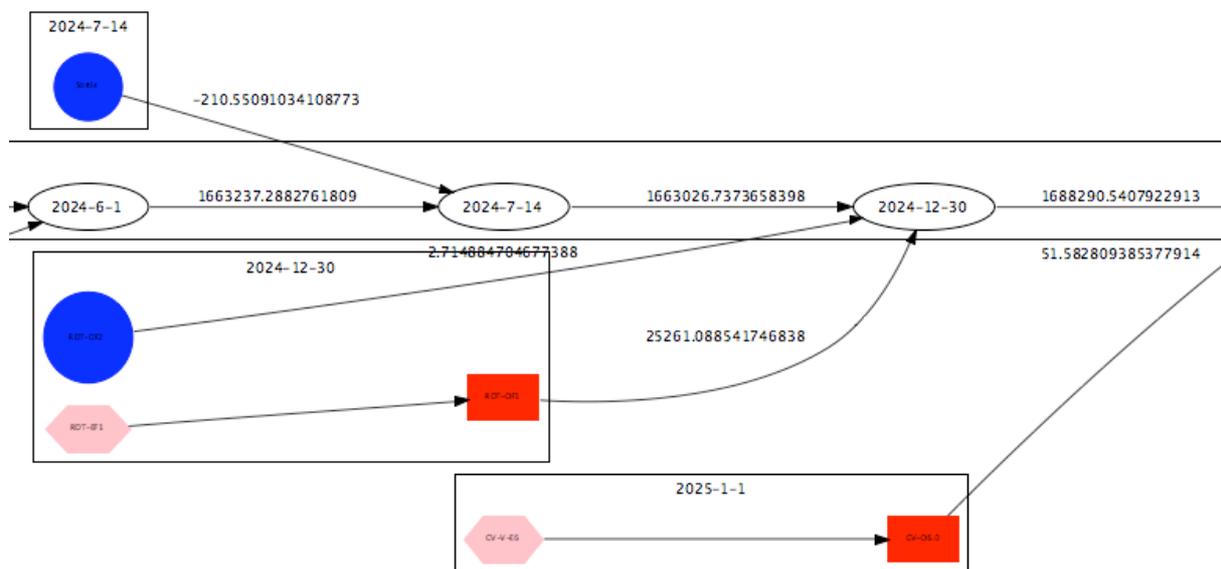


Figure 71 Exemple de graphe produit par Graphviz

Graphviz impose deux dispositifs principaux liés à la représentation de type graphe : les nœuds et les arcs. Le contrat d'écriture dans l'espace de présentation est contraint par ces dispositifs qui sont liés à la nature de la visualisation. Pour chacun de ces dispositifs, il est possible de personnaliser la taille, la valeur, la couleur et la forme. L'utilisation du plan est presque totalement réservée à l'algorithme de génération de graphe. Le format SVG intègre nativement le zoom et le déplacement dans la visualisation produite. Ces moyens d'interaction peuvent être complétés par la définition de liens hypertextes ou de mécanismes beaucoup plus évolués, mais qui nécessiteront de la programmation.

Les deux exemples que nous venons d'aborder illustrent les notions d'espace de représentation, moyens de visualisation et d'interaction, ainsi que les dispositifs. Les

dispositifs influencent largement le contrat d'écriture dans le document que l'on tente de produire. Nous proposons d'aborder plus en avant cette notion dans la section suivante.

Tableau 4 Synthèse d'analyse du format PDF et SVG

Architexte	Espace de représentation	Moyens de visualisation	Variables visuelles exploitables	Moyens d'interaction	Dispositifs
PDF (Acrobat Reader...)	Signets	Visualisation arborescente d'informations textuelles		Liens hypertextes Emboitement des signets	Le Signet
	Document	Visualisation de textes et d'images dans les dimensions du plan du document	Plan Taille Valeur Couleur Orientation Forme	Lien hypertexte Eventuellement annotation	A définir
SVG via Graphviz		Visualisation de graphe dans un plan illimité	Taille Valeur Couleur Forme	Zoom Déplacement (pan) Lien hypertexte Illimité via programmation	Nœud Arc Cluster de Nœuds

### 2.5.c Définir ou interfacier des dispositifs de visualisation et d'interaction

Les capacités de visualisation – qui correspondent aux capacités d'écriture du logiciel auquel on s'intéresse - peuvent être combinées pour former ce que nous appellerons des dispositifs de visualisation.

#### *Notion de dispositif*

Un dispositif de visualisation renvoie tout aussi bien à la définition d'un style textuel – un style correspondant à un ensemble de règles de mise en forme appliquées à un segment textuel – qu'à la manière de dessiner une information dans un diagramme.

Dans notre approche, les dispositifs constituent l'unité élémentaire, la granularité, de l'écriture dans un espace de représentation particulier. Ils sont au cœur de la définition du langage de

visualisation et d'interaction. Dans un espace de représentation particulier, on pourra avoir à disposition différents dispositifs. Dans la zone de signets d'un document pdf, il existe uniquement un type de dispositif : le signet. Le contrat d'écriture permet d'ordonner hiérarchiquement un ensemble de signets. Dans la génération de graphe, on identifie deux dispositifs principaux : les nœuds et les arcs. Le contrat d'écriture dans le graphe permet de lier deux nœuds au moyen d'un arc.

Un dispositif doit être vu comme une configuration particulière d'un ensemble de moyens de visualisation et d'interaction. La Figure 72 illustre le cas d'un dispositif « Nœud » dans le cadre de la génération de graphes. Lorsqu'un nœud est défini dans Graphviz, il va être défini par une couleur, une forme, une taille et une légende particulières. Ces différents éléments constituent les variables visuelles modifiables du dispositif, qui peuvent être instanciées par un éventail plus ou moins large de valeur. Pour le cas particulier du nœud, la forme ou la couleur sont limitées à un ensemble de valeurs, la longueur de la variable est alors contrainte. Par contre, la légende peut contenir n'importe quel texte, et peut être considérée comme une variable dont la longueur est illimitée. Sur l'exemple présenté, les nœuds prennent une forme « invhouse » de couleur verte ou bleue avec une légende particulière pour chacun.

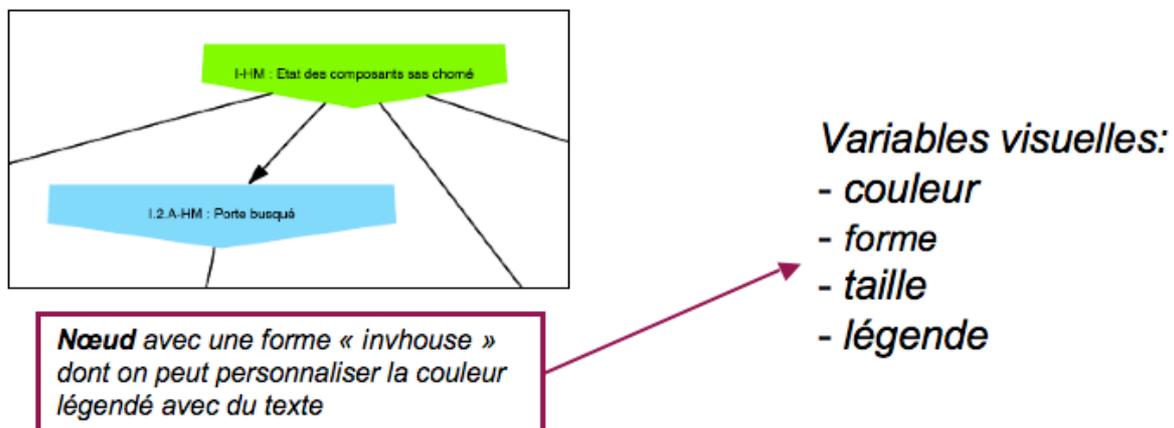


Figure 72 Exemple de dispositif prédéterminé : le Nœud dans Graphviz

D'autres exemples de dispositif, dans le cas de la zone de document dans le format pdf, peuvent être un titre ou un tableau (cf. Figure 73). Ainsi, pour le titre, il sera possible d'en modifier la couleur, la forme correspondant pour le texte, par exemple, à la police d'écriture, à l'italique ou la mise en gras, la taille et bien évidemment le contenu du titre.

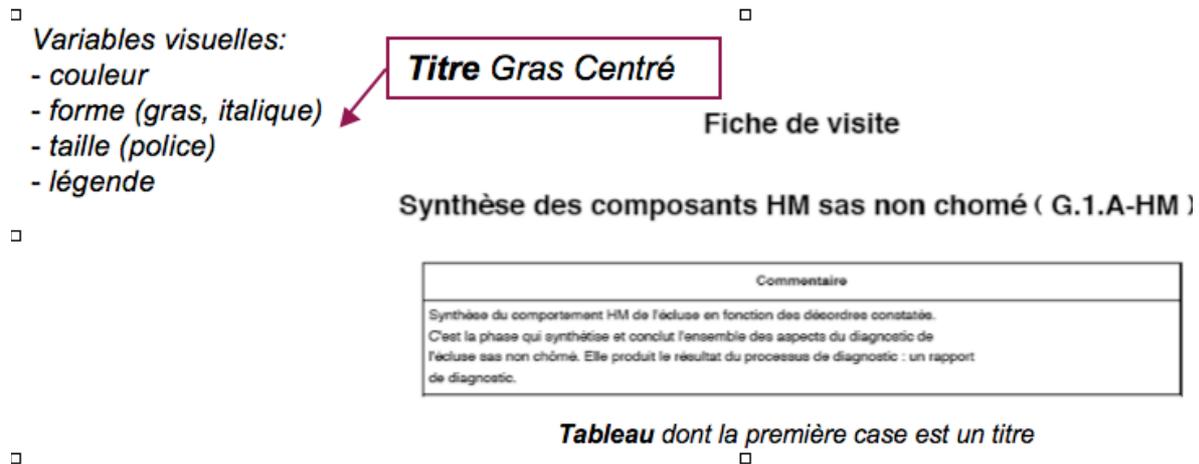


Figure 73 Exemples de dispositifs personnalisés dans un format pdf : le titre et le tableau

### *Types de dispositifs*

Ces dispositifs qui sont en relation avec l'identification et la mise en forme de fragments de texte. Ceci nous amène à dissocier deux types de dispositifs, liés au fait que l'on est confronté à un architecte plus ou moins contraignant :

- Prédéterminé,
- Personnalisé.

Les *dispositifs prédéterminés* peuvent être illustrés par ceux évoqués à l'instant. Ils sont imposés par l'architecte choisi et lié au contrat d'écriture dans les espaces de représentation. Ils représentent une unité de représentation dont la configuration de variables est figée.

Les *dispositifs personnalisés* sont construits de toute pièce par l'auteur du document. La configuration de variables qu'ils traduisent n'est alors pas imposée par l'architecte. L'auteur du document va définir ses propres dispositifs en autonomie, qui traduiront le contrat d'écriture et de lecture. C'est par exemple le cas du titre et du tableau présentés sur la Figure 73.

Les dispositifs personnalisés correspondent aux styles définis dans un traitement de texte Wysiwyg comme Microsoft Word (cf. Figure 74). En effet, l'application d'un style sur une zone de texte dans un document va lui appliquer une mise en forme particulière. Cette mise en valeur structure le document et agit comme un balisage généraliste, puisqu'il délimite le fragment textuel avec un style traduisant une sémantique particulière liée à une intention de

l'auteur (cf. Chapitre 3 -3.2. page 78) : cette opération constitue un balisage généraliste défini de manière autonome par l'auteur.

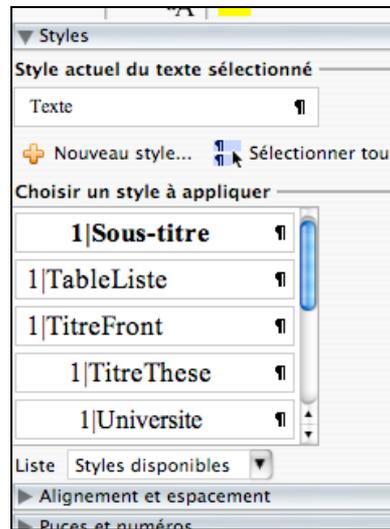


Figure 74 La fenêtre d'application de style dans Microsoft Word.

La définition de dispositifs personnalisés pourra nécessiter une certaine expertise lorsque l'on s'intéressera à des formats effectivement très ouverts mais demandant une compétence, par exemple en programmation, ce qui est souvent l'apanage des informaticiens. Il s'agira alors pour un auteur en collaboration avec l'informaticien de spécifier des bibliothèques de dispositif en relation avec des bibliothèques logicielles existantes, définissant un contrat d'écriture et de lecture dans le document. L'auteur perdra alors une part d'autonomie au regard de la définition de la forme.

#### *Catégoriser les variables des dispositifs*

Un dispositif correspond à une configuration particulière de moyens de visualisation ou d'interaction. Dans sa définition, nous proposons de caractériser ces moyens selon qu'ils sont adaptés ou non pour faire transiter un type de composante. Cette caractérisation va se faire en s'appuyant sur le tableau mettant en relation variables visuelles et niveau de perception (cf. Chapitre 5 -2.2.a page 120) ainsi qu'en considérant les catégories que nous avons ajoutées : « Identifiant » et « Description ».

Ainsi, si l'on reprend l'exemple du dispositif de type nœud présenté sur la Figure 72, les variables peuvent être catégorisées sous la forme du Tableau 5. La catégorisation des variables visuelles de type couleur, forme et taille sont conformes aux propositions de Bertin au regard des catégories Qualitative, Ordonné et Quantitative.

Pour la catégorie Description, nous ne préconisons que de pouvoir utiliser la Légende du dispositif. Les descriptions doivent faire l'objet d'une lecture, qu'il s'agisse d'un texte ou une image, et doivent apparaître sous leur forme originale. On peut noter que s'il s'agit d'un texte et que celui-ci est long, on pourra éventuellement n'en présenter que le début dans le dispositif nœud et proposer d'accéder au contenu intégral de manière interactive.

Les informations de type Identifiant vont pouvoir clairement apparaître dans la légende du nœud. Nous proposons également qu'elles puissent être représentées via la couleur, la forme et la taille mais à condition que leur longueur ne soit pas trop grande. Ceci signifie que le nombre d'identifiants différents à représenter ne doit pas être trop grand.

Tableau 5 Spécification du dispositif « nœud »

<b>Variable</b>	<b>Catégories Sémiotiques</b>				
	<i>Identifiant</i>	<i>Description</i>	<i>Qualitative</i>	<i>Ordonné</i>	<i>Quantitative</i>
<i>Légende</i>	X	X	X	X	X
<i>Couleur</i>	Eventuellement		X		
<i>Forme</i>	Eventuellement		X		
<i>Taille</i>	Eventuellement		X	X	X

## 2.6. Construction du document

Dans notre approche, la construction du document passe par une mise en relation du modèle de perception que nous avons décrit dans la section 2.4. page 189 avec le modèle d'architexte évoqué dans la section 2.5. page 193. Il s'agit dans cette mise en relation de traduire la sémantique des informations à restituer au travers des capacités de l'architexte exploité.

Comme nous l'avons introduit, un architexte peut intégrer plusieurs espaces de représentation, chacun constituant par exemple un mode de visualisation différent sur les informations. Dans la construction du document, il est nécessaire de développer ce que nous nommerons un modèle éditorial de document. Celui-ci précise les espaces de représentation que l'on souhaite mobiliser.

Nous n'approfondirons pas cet aspect ici, bien que nous ayons commencé à l'investiguer. Nous présentons donc la construction du document, partant de l'hypothèse que l'on est face à un seul espace de représentation.

### 2.6.a Du modèle de domaine à l'architexte

Le modèle de perception se positionne en pivot dans le passage du modèle de domaine à l'architexte. La définition du modèle de perception, présentée au 2.4. page 189, abstrait les données issues d'une problématique particulière vers ce que nous avons nommé des catégories sémiotiques. Ces catégories traduisent la manière dont les données doivent être perçues par un utilisateur. Cette tâche de catégorisation relève d'un traitement manuel supervisé.

À la suite de cette catégorisation, on peut associer à une donnée qui constituera une composante de l'information que l'on souhaite visualiser, une ou plusieurs variables visuelles adaptées au niveau de perception auquel elle correspond. Dans cette mise en relation du modèle de perception avec les possibilités de l'architexte, la tâche peut être automatique dans certain cas, mais introduira souvent une part de supervision (cf. 2.6.b page 209). Nous avons évoqué qu'une difficulté était que, dans un contexte réel, il pouvait être difficile d'avoir accès aux variables visuelles pures. Nous introduisons ainsi la notion de dispositif, qui constitue une combinaison particulière de variables visuelles (cf. 2.5.c page 199). Ainsi, le passage du modèle de perception à l'architexte consiste à associer les informations que l'on doit représenter avec un dispositif et les composantes de chaque information à une variable visuelle composant ce dispositif.

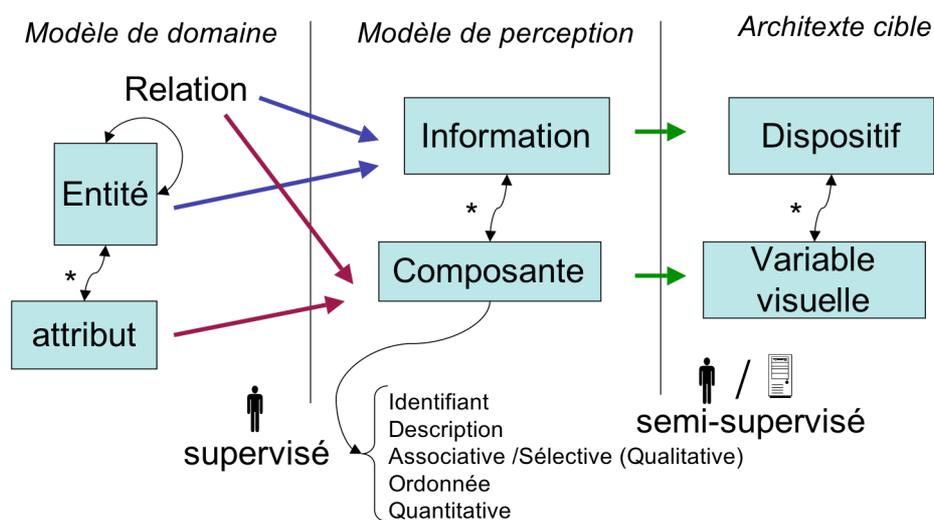


Figure 75 Du modèle de domaine à l'architexte

*Instanciation de dispositif : le cas des guides de diagnostic*

Pour illustrer l’instanciation d’un dispositif, nous reprenons l’exemple des guides de diagnostic. Le modèle de perception correspondant est synthétisé dans le Tableau 3 page 193. Nous allons procéder à la construction d’une représentation graphique du guide en exploitant le dispositif « nœud » dans la section précédente. L’objet est de construire un graphe qui va traduire le document structuré matérialisant le guide de diagnostic. Nous nous limitons ici à décrire l’instanciation des nœuds. Les liens dans le graphe traduisent l’imbrication des éléments structuraux du document ou l’enchaînement du diagnostic (cf. Figure 76).

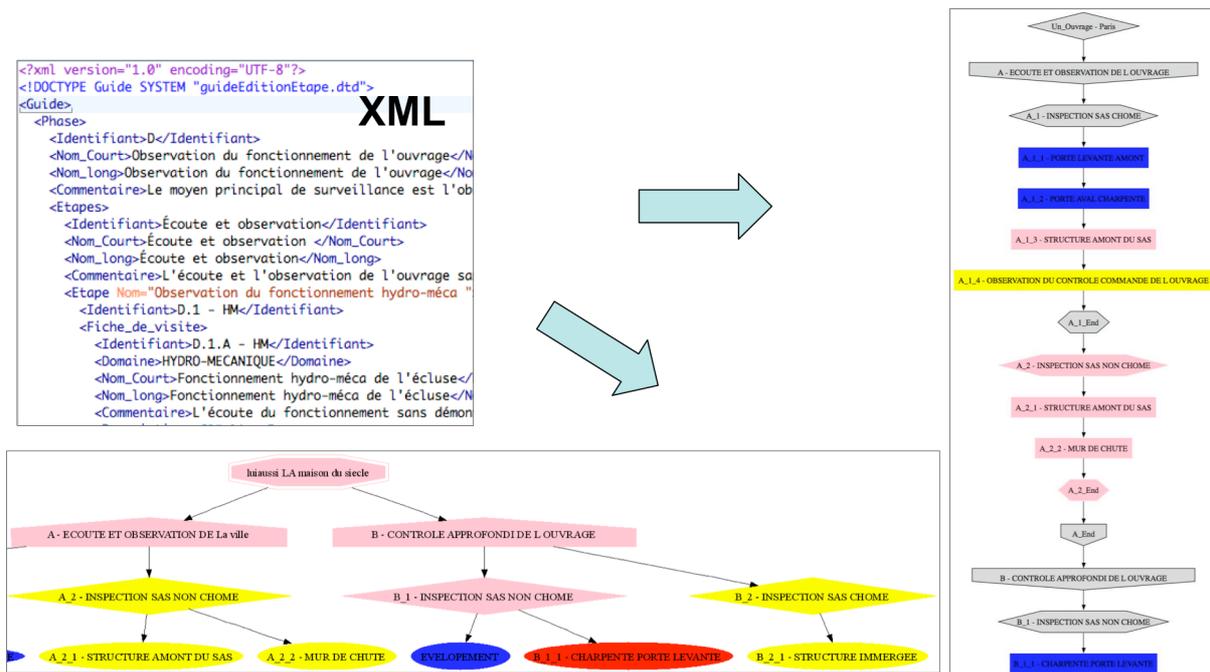


Figure 76 Transformation du document structuré correspondant au guide de diagnostic vers deux visualisations graphiques.

Dans la réalisation de ces graphes, nous allons représenter nos informations au travers de nœuds. Les informations correspondent aux éléments structuraux qui définissent le découpage séquentiel du diagnostic, c’est à dire : Phase, Etapes, Etape et Fiche de visite. Les composantes de ces informations doivent ensuite être associées aux variables visuelles paramétrables sur le dispositif : la légende, la couleur, la forme, et la taille.

Le « Nom long » et « l’Identifiant » sont caractérisés comme Identifiant. Ils constitueront la légende du nœud. On procèdera à une concaténation des deux contenus.

Le « Domaine » ainsi que les noms de balises correspondent à deux composantes qualitatives. Forme et couleur sont adaptées pour faire transiter cette catégorie (cf. Tableau 5 page 203).

Le type d'élément correspondant à la balise pourra être traduit par une forme, alors que le domaine sera rendu visible par la couleur (cf. Figure 77).

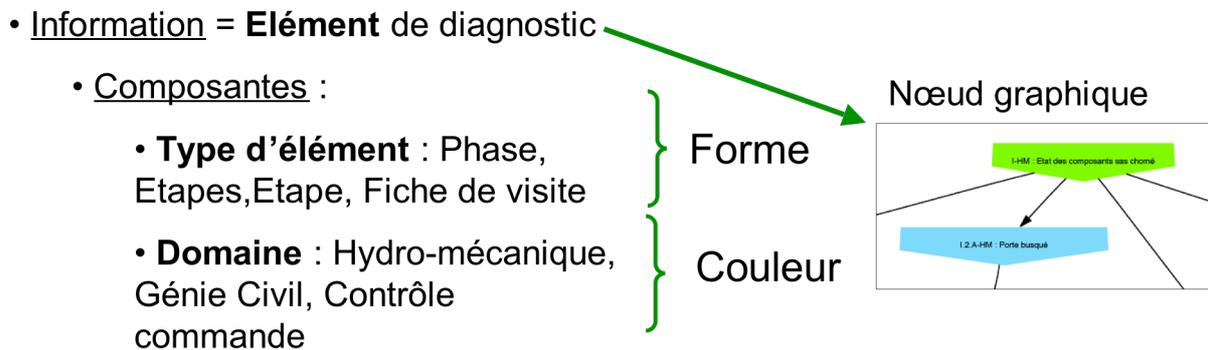


Figure 77 Mise en relation d'une information avec un dispositif

Pour terminer la définition de la sémantique de construction du document, qui correspond ici à une représentation graphique, il est nécessaire d'étudier les contenus délimités par les éléments structuraux. Ainsi, on doit identifier, pour chaque composante qualitative, le jeu de valeurs qu'elle peut prendre. Dans notre exemple, on peut dénombrer quatre types d'éléments différents : Phase, Etapes, Etape, Fiche de visite. On peut également dénombrer trois domaines : Hydro-mécanique, Génie Civil et Contrôle Commande. Pour chaque valeur on associera une valeur de variable visuelle disponible. On aboutit ainsi au résultat présenté sur la Figure 78.

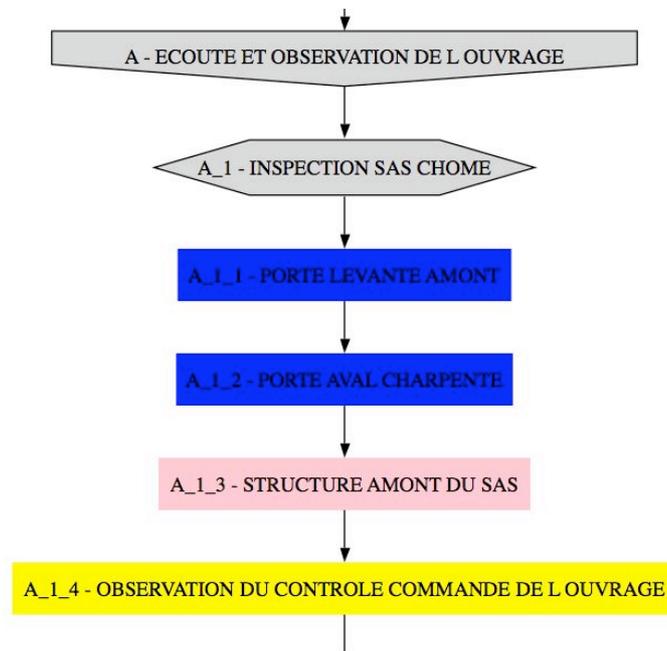


Figure 78 Exemple de visualisation graphique du guide de diagnostic

La sémantique visuelle de construction du document peut alors être synthétisée comme sur la Figure 79. Il s'agit d'une représentation du langage inhérent au document produit selon l'axe syntagmatique et paradigmatique (Barthès 1982) qui constitue le contrat d'écriture et de lecture dans le document.

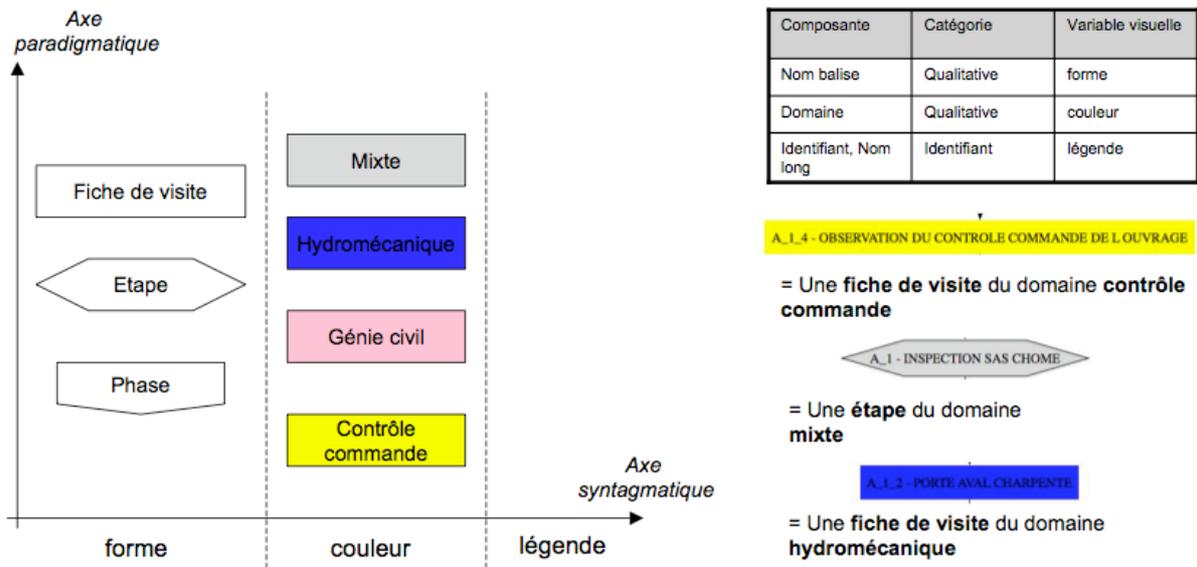


Figure 79 Sémantique visuelle de construction du document

Ce type d'approche permet de construire des restitutions du même jeu de données en entrée à travers d'autres architectures. Ainsi, le Tableau 6 illustre la mise en relation du guide avec un lecteur de fichier PDF pour aboutir au résultat présenté sur la Figure 80. Une autre illustration pour le cas d'un navigateur WEB et du format HTML est proposée dans le Tableau 7 pour aboutir à un document WEB présenté sur la Figure 81.

Tableau 6 Définition d'un système de visualisation dans le cas du PDF

Informations / Composantes	Dispositif ou variable dans la zone document du PDF	Dispositif ou variable dans la zone de signets du PDF
Phase, Etapes, Etape, Fiche de visite	Titre et saut de page	Signet
Identifiant, Nom long	Légende	Légende

Description	Cadre textuel	
Nom élément	Légende, Taille (Police)	
Domaine	Légende	

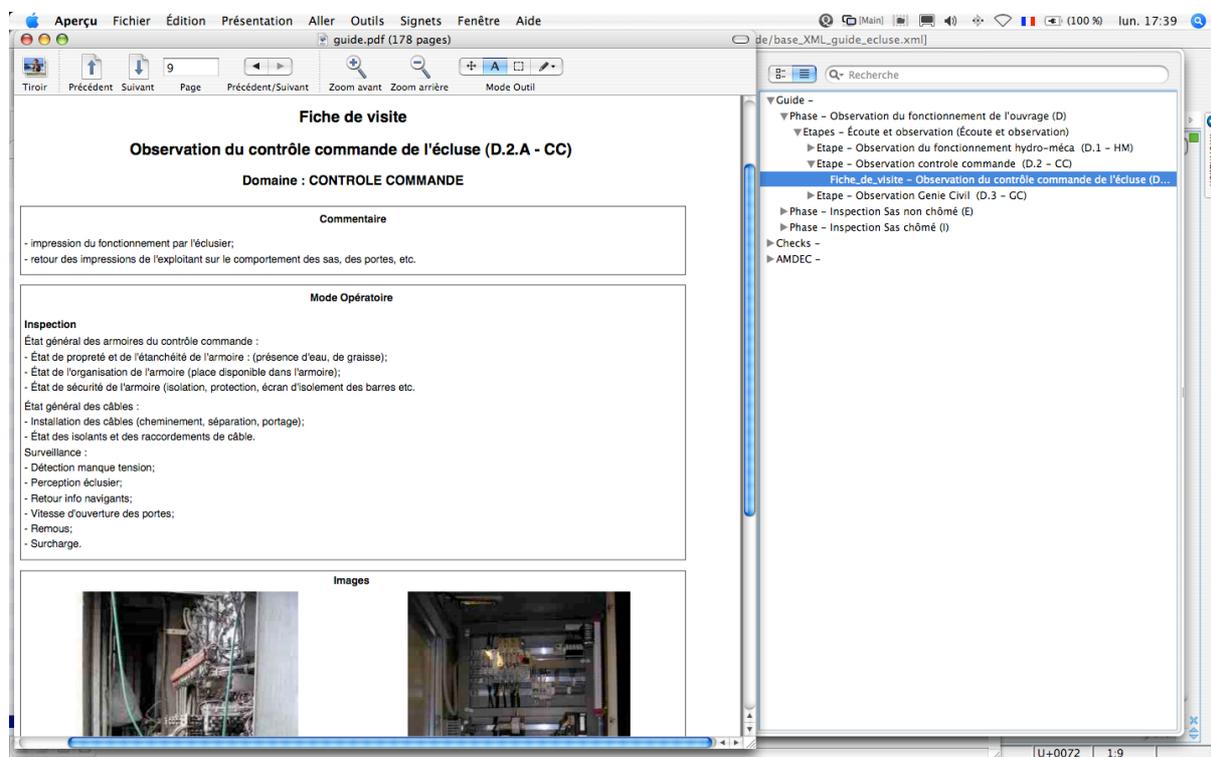


Figure 80 Le guide sous la forme d'un fichier PDF adapté à une lecture sur écran et à l'impression

Tableau 7 Définition d'un système de visualisation dans le cas du format HTML

Informations / Composantes	Dispositif ou variable dans la zone document du HTML	Dispositif ou variable dans la zone d'arbre de navigation du HTML
Phase, Etapes, Etape, Fiche de visite	Titre et saut de page	Noeud
Identifiant, Nom long	Légende	Légende
Description	Cadre textuel	
Nom élément	Couleur, Légende, Taille (Police)	Couleur, Légende
Domaine	Légende	

**- Ph - D - Observation du fonctionnement de l'ouvrage -**

**Commentaire :**  
Le moyen principal de surveillance est l'observation visuelle et auditive. Il est par conséquent indispensable de lui donner la régularité, la rigueur et la précision qui s'imposent. Cette phase d'inspection est réalisée par l'exploitant

**- Es - Ecoute\_et\_observation - Ecoute et observation -**

**Commentaire :**  
L'écoute et l'observation de l'ouvrage sans démontage dans les domaines :  
- Hydro-mécanique  
- Génie civil  
- Contrôle commande

**- E - Observation du fonctionnement hydro-méca - D.1 - HM -**

**- FdV - D.1.A - HM - Fonctionnement hydro-méca de l'écluse -**

**Domaine :** HYDRO-MECANIQUE

**Commentaire :**  
L'écoute du fonctionnement sans démontage :  
- Placé à proximité des organes en mouvement (système de manoeuvre des portes, des vannes aqueduc, accouplement, palier), il s'agit d'écouter et observer les composants en fonctionnement comme par exemple la chaîne cinématique de manoeuvre des organes etc.

**- Description -**

Figure 81 Le guide sous la forme d'une page HTML

### 2.6.b Approche supervisée vs approche non supervisée

Comme nous venons de le voir, la méthode de restitution proposée aboutit à la définition d'un langage qui traduit la sémantique des données que l'on souhaite restituer dans les capacités de visualisation et d'interaction inhérente à l'interface de sortie pour laquelle on a opté. La définition de ce langage passe par une analyse de la nature des données à restituer pour s'écarter de la sémantique métier afin de monter en abstraction à un niveau perceptif plus proche de la visualisation. Cette tâche relève d'une interprétation humaine pour aboutir à la définition du modèle de perception.

D'un autre côté, il faut procéder à une analyse de l'architexte choisi afin d'identifier et de qualifier ses possibilités. Le modèle d'architexte permet ainsi d'évaluer si une variable visuelle ou interactive est adaptée ou non pour faire transiter tel ou tel type de composante. Là encore cette tâche tient à une opération supervisée en relation avec une interprétation humaine.

Le cœur de la construction du langage dont nous traitons ici est la mise en relation du modèle de perception avec le modèle d'architexte. Nous en avons décrit de manière synthétique les mécanismes dans la section précédente en nous appuyant sur un exemple. La méthode évoquée laisse entrevoir la possibilité que cette mise en relation soit faite :

- de manière manuelle,

- éventuellement de manière automatique.
- de manière semi-supervisée,

La *méthode manuelle* est en fait conforme à l'explication que nous avons pu faire de la mise en relation du modèle de perception avec le modèle d'architexte.

La *méthode automatique* se traduirait par le développement d'un algorithme qui va procéder à l'affectation de dispositifs aux informations puis des variables de ces dispositifs aux composantes en tenant compte l'adéquation entre le type de composante et la variable. Une ébauche de ce travail a été faite dans le cadre d'un projet proposé à des étudiants participant à une unité d'enseignement en programmation objet. Les résultats produits ont permis de montrer la faisabilité de ce type de développement sachant qu'il s'agirait d'en évaluer le passage à l'échelle réelle mais également la pertinence de la production d'un document de manière automatique (cf. Chapitre 3 -2.2. page 76).

La *méthode semi-supervisée* se situe à mi-chemin ou plus exactement à l'intersection entre une méthode intégralement manuelle et une méthode automatique. L'idée est alors de proposer un processus interactif permettant de co-construire le langage de visualisation avec la machine à partir d'un modèle de perception et d'un modèle d'architexte déjà spécifié. Ce processus interactif passera alors en revue successivement les informations et les composantes à restituer, en proposant les dispositifs et variables visuelles qu'il est possible de leur associer. Par exemple, pour chaque valeur possible d'une composante qualitative, on pourra proposer à un utilisateur une palette de couleurs.

La définition du modèle de perception pourra également être appuyée par la machine. Il s'agira par exemple à partir d'un document structuré en entrée d'en extraire le modèle de document correspondant. Ce modèle de document recense à la fois les éléments structuraux exploités et leur organisation (cf. Chapitre 3 -1.3.c page 69). Dans un processus interactif on pourra alors proposer à l'utilisateur de catégoriser les différents éléments structuraux qui apparaissent dans le document.

La définition des modèles d'architextes doit être opérés en amont de ce processus interactif. Il s'agit dans un premier temps, d'identifier les architextes avec lesquels on souhaite être en mesure de travailler. Dans un deuxième temps, il convient de définir des bibliothèques de dispositifs en spécifiant les variables disponibles et leur adéquation avec une ou plusieurs catégories sémiotique. La définition des modèles d'architexte constituent un travail en amont

impliquant en particulier une expertise en informatique dans la définition des bibliothèques de dispositifs.

### 3 - Vision intégrée de la capitalisation et de la restitution sur la base de documents structurés

Dans ce chapitre, nous avons décrit une proposition d'approche à la fois méthodologique et instrumentale, basée sur l'utilisation de documents structurés, permettant d'appuyer l'activité de capitalisation et de restitution, dans une démarche de gestion des connaissances. Bien que ces deux activités aient été présentées de manière indépendantes dans ce chapitre, il est possible de définir une vision intégrée de la démarche de gestion des connaissances en s'appuyant sur notre proposition.

Sans volonté de fournir une solution « tout en un » de gestion des connaissances, et toujours en s'assurant de pouvoir être en adéquation avec les infrastructures existantes, il semble possible de spécifier ce que nous nommerons un *Atelier Sémiotique*. Cet atelier se positionnerait alors comme un carrefour dans la démarche de capitalisation des connaissances, dévouée à la restitution de document structuré. Nous le qualifions de sémiotique, au sens où le processus de restitution, basé sur l'approche que nous avons décrite, permettrait de traduire le contenu de documents structurés vers les capacités de visualisation et d'interaction d'un architecte cible, en conservant la sémantique de la traduction. L'atelier s'intégrerait alors dans la démarche de gestion des connaissances comme un instrument permettant par exemple d'initier la capitalisation dans un format tout en ayant la possibilité de passer à un autre plus adapté au fur et à mesure de l'avancée dans la capitalisation.

L'objectif sous-jacent est d'essayer de fournir un outil ne nécessitant pas une expertise informatique trop importante. Il s'agit de rendre la restitution d'un document structuré accessible à des non-informaticiens en s'appuyant sur une plate-forme intégrant :

- Des bibliothèques de dispositifs de visualisation et d'interaction
- Un moteur de restitution documentaire
- Un accompagnement à la construction des documents

Les *bibliothèques de dispositifs* doivent permettre de fournir à l'utilisateur un ensemble de moyens permettant de restituer le document structuré dans différents systèmes de

représentation et adapté à différents formats ou architextes cibles. Ces bibliothèques doivent être produites en amont et constituer une ressource valorisable dans différents projets.

Le *moteur de restitution* documentaire doit être développé en intégrant la philosophie de l'approche de restitution que nous avons présentée dans la section 2 - page 180. Ce moteur a pour objet de transformer un document structuré en traduisant les briques élémentaires qui le constituent dans une forme particulière basée, sur l'utilisation des bibliothèques de dispositifs. La construction du document pourrait alors s'appuyer sur une logique, supervisée ou non, telle que l'avons évoqué au 2.6.b page 209.

Nous avons évoqué au-dessus l'importance d'offrir un outil adapté à des non-informaticiens. C'est donc à ce niveau qu'il est important d'intégrer *un accompagnement à la construction des documents* lorsque celle-ci est supervisée. Dans la section 2.6.b page 209, nous avons décrit la possibilité de développer une solution interactive à destination des utilisateurs permettant sous la forme d'un dialogue de construire le document en interaction avec le système. Tout d'abord, il s'agirait de développer un ensemble d'interface permettant à l'utilisateur de classifier les éléments structuraux constituant le document structuré. Ceci permettrait de générer le modèle de perception. À partir du modèle de perception, le système pourrait proposer un ensemble de moyens de visualisation et d'interaction jugés comme adaptés conformément aux modèles d'architextes prédéfinis.

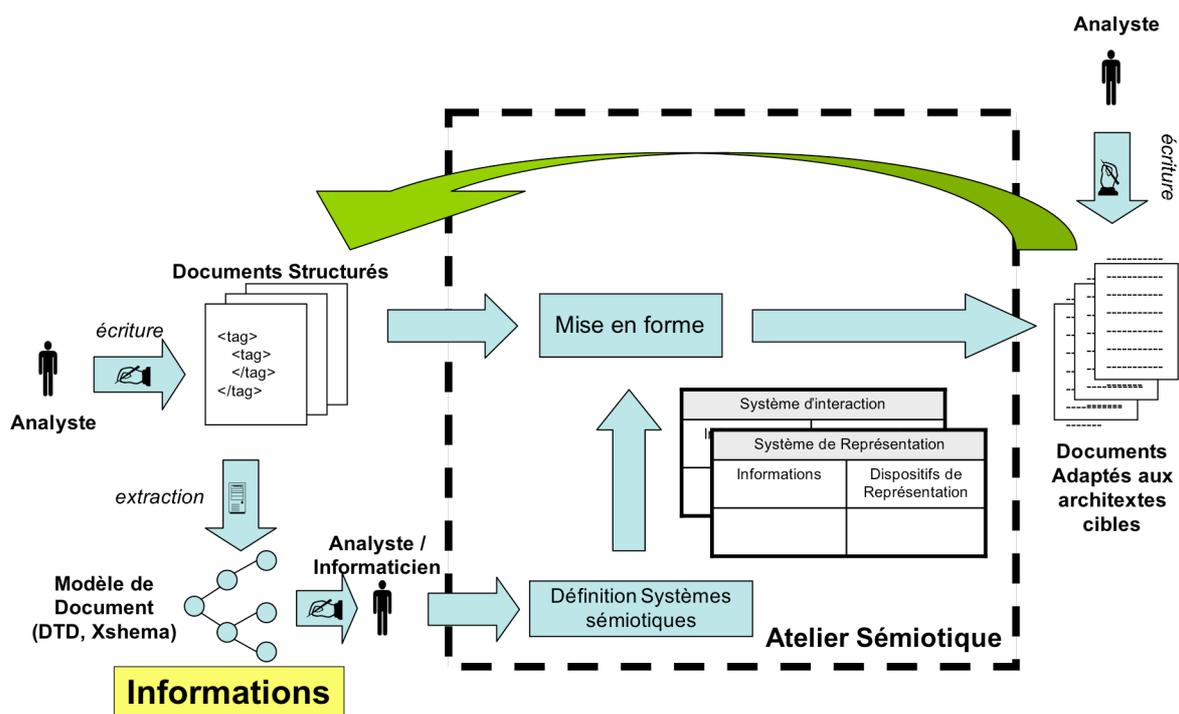


Figure 82 Vision intégrée de la démarche de gestion des connaissances appuyée par un atelier  
sémiotique

# Chapitre 11 - Développement et mise en œuvre

## 1 - Vision globale de la mise en œuvre

L'approche que nous avons décrite dans le chapitre précédent, si elle introduit une dimension méthodologique importante, s'appuie également sur un outillage important des activités et une adaptation au contexte réel. Notre approche a été construite au travers d'une étude empirique de « recherche - action », qui nous a permis de développer un certain nombre de ressources et d'outils ainsi que de pouvoir proposer aujourd'hui un scénario instrumenté de capitalisation et restitution de connaissances capitalisées (cf. Figure 83).

Nous pouvons dissocier 2 contextes de gestion des connaissances dans lesquels peut s'insérer notre approche. En premier lieu, elle représente un élément de solution dans le cadre d'une nouvelle démarche de gestion de connaissances. Il s'agit alors d'accompagner l'ensemble de la démarche allant de la capitalisation à la restitution de connaissances capitalisées.

Elle peut également se porter en complément d'un existant. Par exemple, il s'agira de compléter ou faire évoluer un système à base de connaissances déjà existant pour rendre la capitalisation plus flexible ou encore offrir une souplesse dans la restitution et permettre de définir un éventail de visualisations possibles adaptées à différents architectes.

### 1.1. La capitalisation au travers de documents structurés

L'objet n'est pas ici de décrire l'intérêt du document structuré dans l'activité de capitalisation. Nous avons déjà traité de son utilisation dans cette activité au Chapitre 10 -1.2. page 172. Nous proposons de préciser tout d'abord l'éventail d'outils que nous avons mobilisé dans la conception et la rédaction de documents structurés en mettant l'accent sur leur intérêt. Nous poursuivons par la description de la mise en œuvre du document structuré comme espace de capitalisation face à un système informatique à base de connaissances déjà existant.

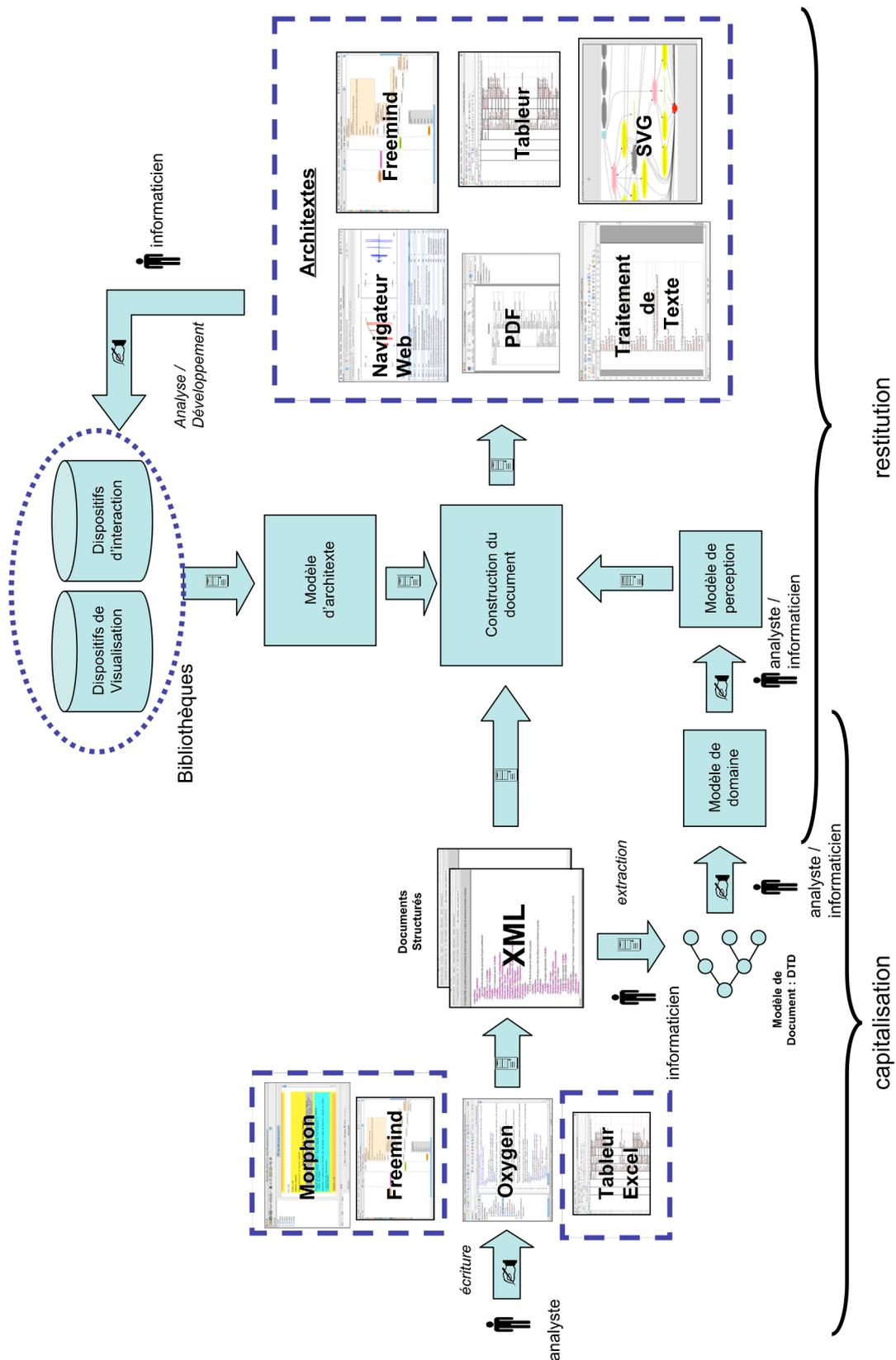


Figure 83 Scénario de capitalisation et restitution dans une action de gestion de connaissances.

### *1.1.a Un éventail d'outils nécessaire*

Notre étude nous permet de définir un éventail d'éditeurs nécessaire pour la conception et la rédaction de documents structurés. Chaque éditeur de document structuré offre ainsi une visualisation adaptée pour traiter un aspect du document structuré. Cet éventail doit alors être composé par :

- Un éditeur de code source
- Un éditeur Wysiwym
- Un éditeur structurel

L'*éditeur de code source* de document structuré est un outil initialement destiné à un public d'informaticiens puisqu'il offre une visualisation de l'information sous la forme d'un code source balisé (cf. Figure 84). Nous avons toutefois pu observer qu'un non-informaticien pouvait en faire une utilisation tout à fait correcte (cf. Chapitre 10 -1.2.a page 173). Ce type d'éditeur permet une saisie en extension des documents, constructiviste, permettant de créer et modifier le contenu et les éléments structuraux. Il introduit de nombreux mécanismes permettant d'accompagner la saisie de documents structurés tel que la complétion du balisage ou encore la vérification en contexte du caractère bien formé ou valide du document (cf. Chapitre 3 -1.3.c page 69). Ce type d'éditeur intègre également en général un mécanisme permettant de générer un modèle de document relatif à un document structuré en cours de rédaction. L'éditeur de code source se révèle donc particulièrement adapté dans la conception initiale de document structuré, à un stade où l'on possède une vue peu précise sur la structure ou encore le contenu. Il semble également adapté lorsque la structure du document est déjà définie. En effet, il est alors possible de faire appel à un modèle de document (cf. Chapitre 3 -1.3.c page 69) qui assistera un utilisateur à se conformer à un mode particulier de structuration. L'éditeur de code source offre alors une aide à la saisie des « documents » qui peut être « documentée ». Par exemple, l'éditeur de code source Oxygen permet de visualiser les éléments de documentation du modèle de document durant la saisie. Durant la rédaction du document, on peut alors obtenir des éléments permettant de mieux identifier le sens d'une balise.

L'*éditeur wysiwym* (« What you see is what you mean ») vient en complément de l'éditeur de code source. Il se destine particulièrement aux utilisateurs non informaticiens puisque l'on constate qu'un éditeur de code source peut poser quelques difficultés lorsque le contenu du

document structuré s'étend. L'abondance des balises qui délimitent le contenu du document peut poser des difficultés de lisibilité par exemple. Cette difficulté peut également être ressentie par un informaticien mais dans une moindre mesure car ils possèdent souvent une maîtrise de l'éditeur leur permettant de mieux se repérer dans le document. Un éditeur wysiwym doit permettre de masquer les éléments structuraux en les traduisant par exemple dans une sémantique visuelle. C'est par exemple le cas de l'éditeur Morphon présenté sur la Figure 85. Sur cet exemple, les balises ont été traduites par des cadres colorés et une mise en forme particulière des contenus délimités. Ce type de solution se révèle adapté pour faire évoluer les contenus des documents structurés. Le masquage des balises se fait au prix de la spécification d'une feuille de style – en l'occurrence ici une feuille de style CSS – qui peut être construite sans une expertise informatique par une interface d'accompagnement intégrée à l'éditeur.

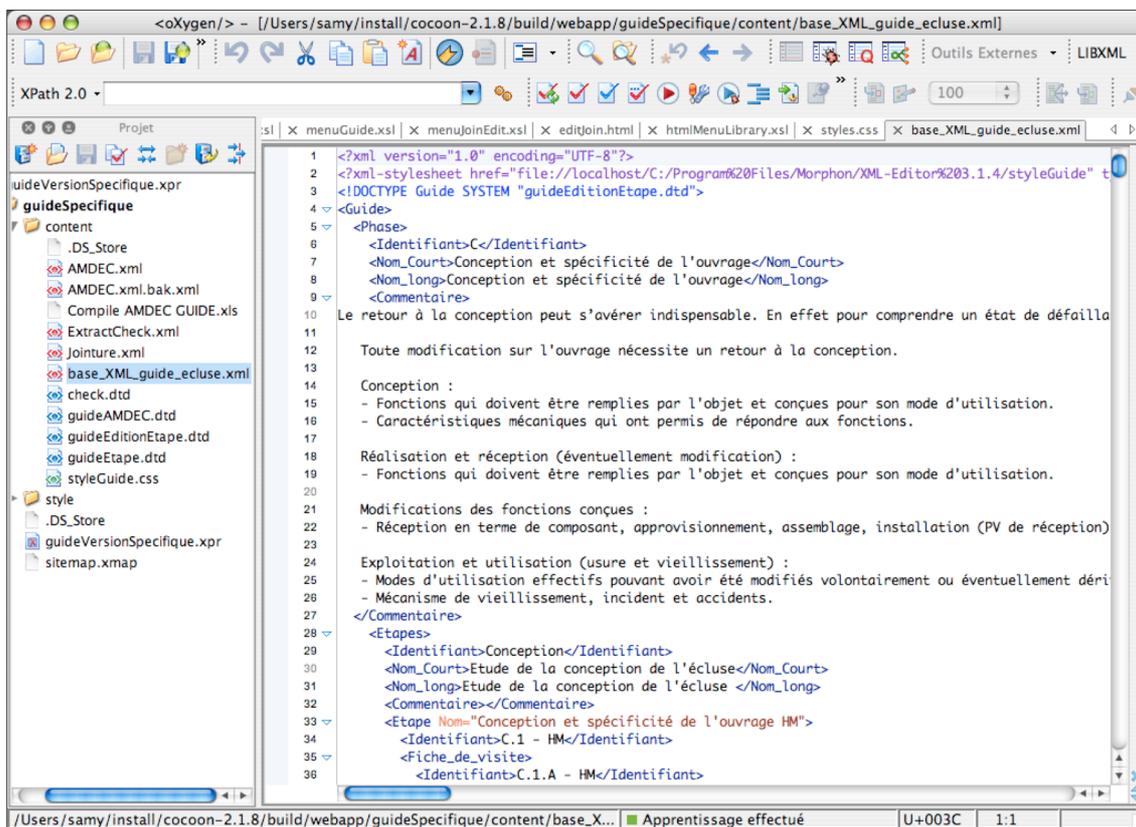


Figure 84 L'éditeur de code source XML Oxygen

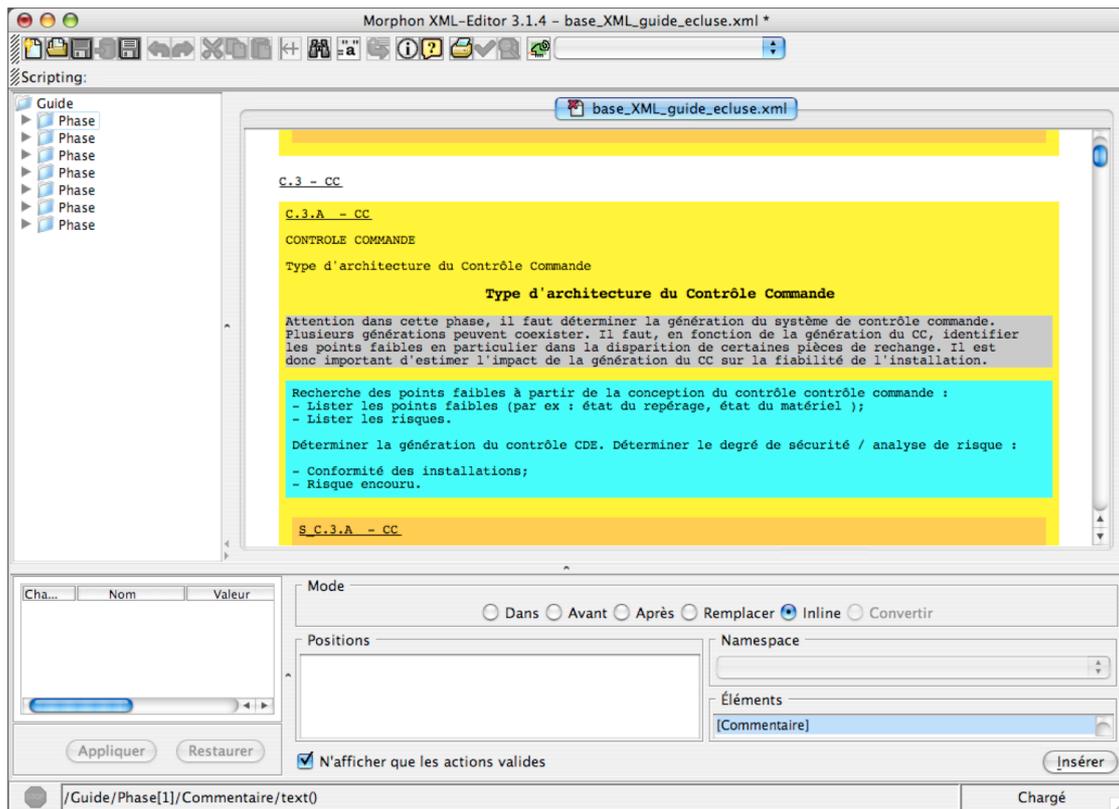


Figure 85 Une vue sur l'interface de saisie proposée par Morphon

L'*éditeur structurel* doit permettre de se focaliser sur la structure hiérarchique inhérente au document que nous développons. Il traduit le document structuré sous la forme d'une arborescence interactive. Les éditeurs de code source évolués, comme par exemple Oxygen, proposent le plus souvent une forme de visualisation du document centrée sur l'édition de l'arborescence. L'exemple présenté sur la Figure 86 tient plus de la liste indentée que de la métaphore d'un arbre. De plus, la visualisation est très épurée, au sens où elle traduit le contenu du document structuré sous la forme de champs de saisie sans mise en forme particulière. Ce type de solution reste donc plutôt préconisé une fois encore à un public d'informaticiens ayant une bonne compréhension du balisage et n'ayant pas besoin d'une sémantique visuelle pour mieux l'appréhender.

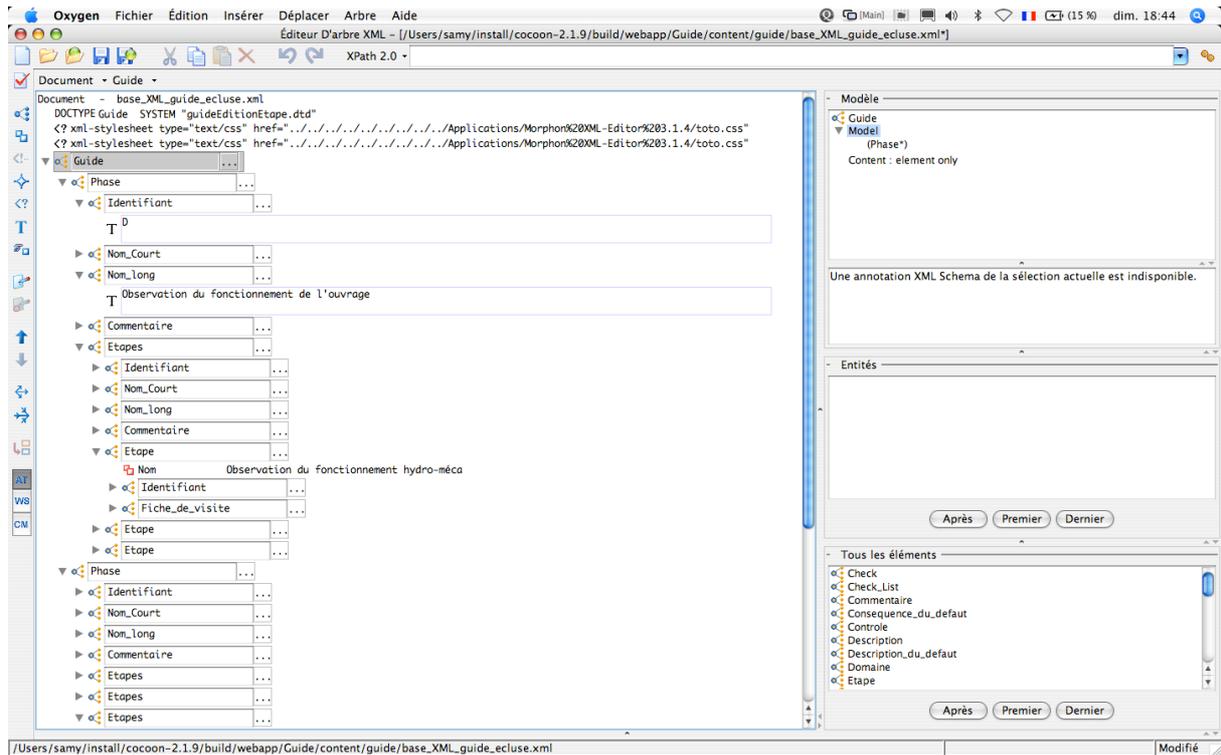


Figure 86 Editeur d'arbre XML dans Oxygen

Dans notre mise en œuvre, notre éditeur structurel correspond en fait à un éditeur de cartes conceptuelles (Freemind). Nous l'avons déjà introduit au Chapitre 10 -1.3.b page 177. Cet éditeur est conçu pour saisir des cartes conceptuelles qui correspondent à une représentation graphique d'informations relationnelles sous la forme d'une arborescence (cf. Figure 87). Il s'agit d'un outil qui propose une prise en main aisée de par sa simplicité. Nous préférons cette solution à l'éditeur d'arbre XML car d'une part la représentation graphique de l'arbre est plus évidente et d'autre part les cartes conceptuelles offrent une marge de manœuvre importante sur la mise en forme des éléments qui la constituent.

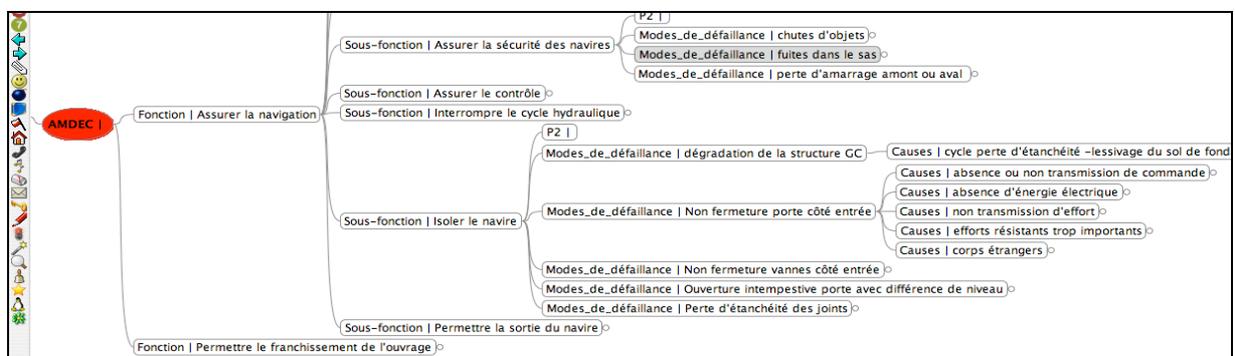


Figure 87 Carte conceptuelle Freemind calculée sur la base d'un fichier XML

Nous proposons également une alternative à l'éditeur de cartes conceptuelles qui résulte de notre étude empirique. Lors de l'analyse de la forme des documents de capitalisation produit (cf. Chapitre 8 -1 - page 148), nous avons souligné que le tableur Excel était assez généralement exploité comme un outil de structuration hiérarchique de données tel que présenté sur la Figure 88. La hiérarchie du document est alors traduite sous une forme matricielle. Ce mode de saisie, qui ne constitue pas une pratique isolée, représente donc une solution qui permet à la fois de traiter le document structuré au niveau de sa structure et de son contenu. Le tableur constitue alors un éditeur structurel tout à fait adapté.

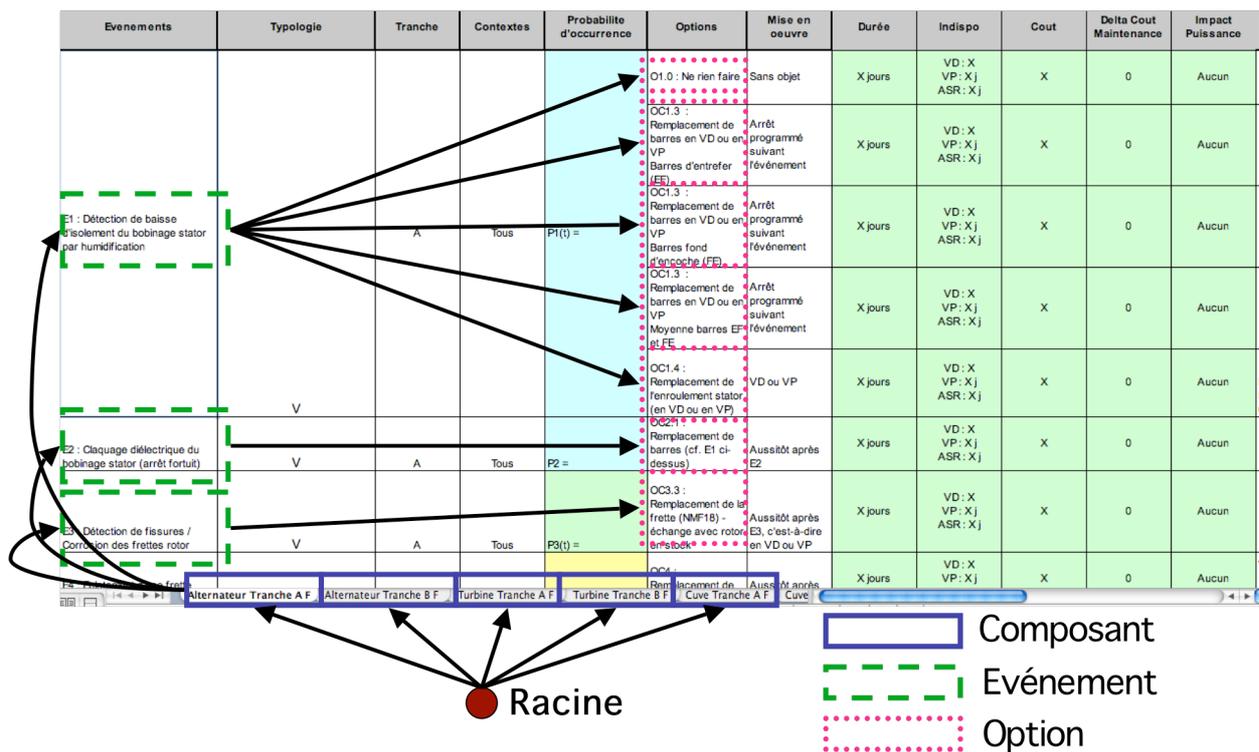


Figure 88 Fichier Excel exploité comme espace de structuration d'information

Nous évoquerons au passage ici l'intégration dans la dernière version de l'éditeur de code source Oxygen d'un mode de saisie similaire présenté sur la Figure 89.

Guide	Phase (3 rows)	Identifiant	Nom_Court	Nom_long	Commentaire	Etapes																									
	1	D	Observation du fonctionnement de l'ouvrage	Observation du fonctionnement de l'ouvrage	Le moyen principal de surveillance est l'observation visuelle et auditive. Il est par conséquent indispensable de lui donner la régularité, la rigueur et la précision qui s'imposent. Cette phase d'inspection est réalisée par l'exploitant	<table border="1"> <thead> <tr> <th>Etapes</th> <th>Identifiant</th> <th>Nom_Court</th> <th>Nom_long</th> <th>Commentaire</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> <td></td> <td></td> <td>Écoute et observation</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td>Écoute et observation</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td>Écoute et observation</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td>L'écoute et l'observation de l'ouvrage sans démontage dans les domaines : - Hydromécanique - Genie civil - Contrôle commande</td> </tr> </tbody> </table>	Etapes	Identifiant	Nom_Court	Nom_long	Commentaire					Écoute et observation					Écoute et observation					Écoute et observation					L'écoute et l'observation de l'ouvrage sans démontage dans les domaines : - Hydromécanique - Genie civil - Contrôle commande
Etapes	Identifiant	Nom_Court	Nom_long	Commentaire																											
				Écoute et observation																											
				Écoute et observation																											
				Écoute et observation																											
				L'écoute et l'observation de l'ouvrage sans démontage dans les domaines : - Hydromécanique - Genie civil - Contrôle commande																											
	2	E	Inspection Sas non chôme	Inspection Sas non chôme	Lors de cette phase, il est nécessaire de rester au moins une journée sur l'ouvrage en situation de fonctionnement, avec un trafic	<table border="1"> <thead> <tr> <th>Etapes (5 rows)</th> </tr> </thead> <tbody> <tr> <td></td> </tr> </tbody> </table>	Etapes (5 rows)																								
Etapes (5 rows)																															

Figure 89 La saisie sous forme de grille d'un fichier XML dans la version 9 de Oxygen

### 1.1.b Face à un système informatique à base de connaissances existant

Les différents outils que nous venons d'évoquer laissent apparaître l'importance de proposer des solutions de conception ou de saisie des documents structurés adaptées aux non informaticiens. Face à un système à base de connaissances existant, il s'agit d'une orientation similaire au sens où il est nécessaire de mettre la saisie des données à la portée de non spécialistes sur le système. Dans notre contexte d'étude, il s'agit également de passer outre les difficultés liées à la mise en œuvre de formulaire (cf. Chapitre 8 -2.1. page 154) afin de profiter des qualités du document structuré dans la saisie d'informations formelle (cf. Chapitre 9 -2.1.c page 164).

La mise en œuvre du document structuré dans cette situation passe donc en premier lieu par le développement d'un modèle de document qui va permettre de cadrer la saisie. Ce modèle de document doit à la fois être construit sur la base des structures de données relatives au système existant ainsi qu'au mode de saisie que l'on souhaite proposer. Le modèle de document doit donc résulter d'une coopération entre un informaticien, qui apporte son expertise dans sa définition, et l'analyste ou encore tout utilisateur qui sera en position de saisie face au système.

Dans la perspective de coupler notre approche et nos outils à un système existant, il est également nécessaire de développer des mécanismes d'importation et d'exportation des documents structurés permettant d'intégrer le contenu du document structuré à un système ou

encore d'obtenir une image d'une base de connaissances existante sous la forme d'un document structuré. L'informaticien reste donc un acteur incontournable dans ce cas de figure.

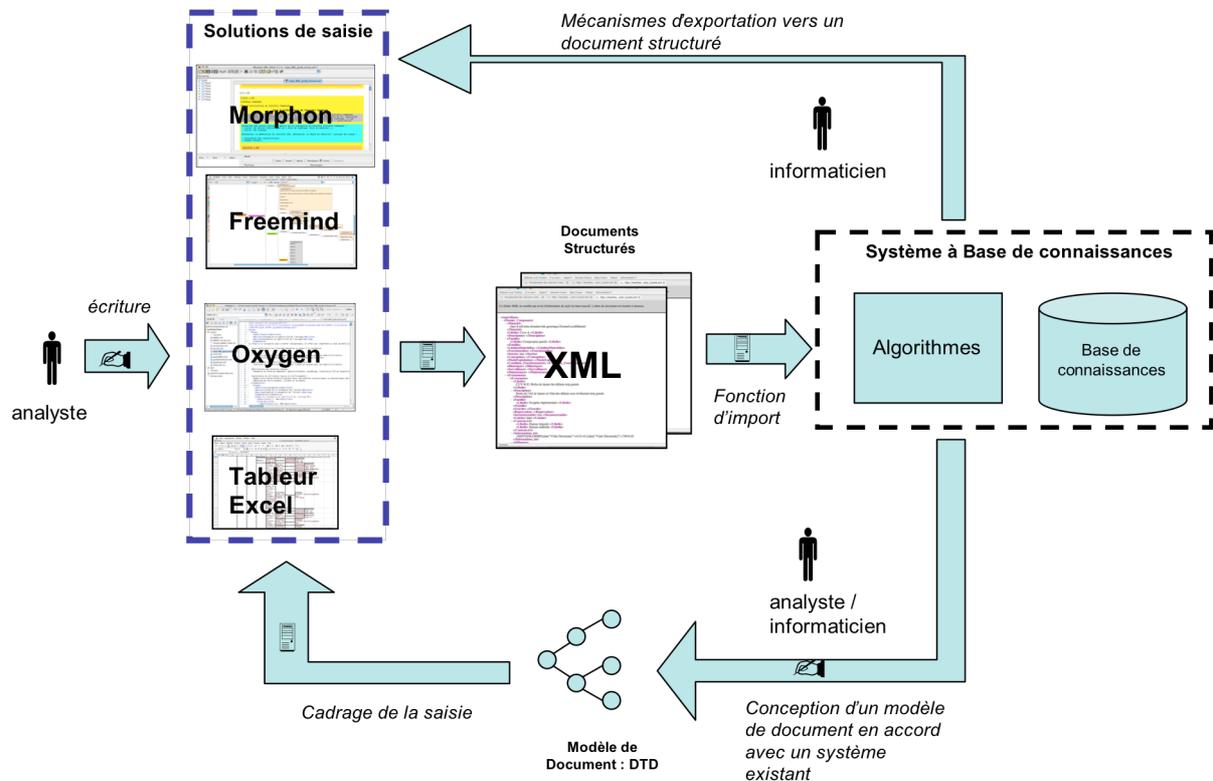


Figure 90 Capitalisation au travers de documents structurés face à un système à base de connaissances existant

L'introduction du document structuré dans le cadre d'informations formelles ne doit pas uniquement en faciliter la saisie par exemple au travers des outils présentés dans la section précédente. Il s'agit de profiter de notre approche de restitution afin de visualiser le document sous différentes formes adaptées à différents architectes.

## 1.2. La restitution

La restitution des documents structurés est fondée sur un moteur de restitution documentaire qui est décrit sur la Figure 89. Ce moteur procède à la construction d'un document mis en forme et adapté à un architecte à partir :

- d'un ou plusieurs documents structurés,
- d'un modèle de perception,

- d'un modèle d'architexte.

Le *document structuré* résulte d'un processus de capitalisation tel que nous avons pu l'évoquer dans la section précédente et constitue la matière à restituer.

Le *modèle de perception* associe les éléments structuraux contenus dans le document structuré à un niveau perceptif. Il s'agit de se détacher de la sémantique métier des informations pour se focaliser sur la manière dont elles doivent être perçues dans la visualisation. La définition de ce modèle de perception a été décrite au Chapitre 10 -2.4. page 189.

Enfin, le *modèle d'architexte* (cf. Chapitre 10 -2.5. page 193) recense les moyens de visualisation et d'interaction d'un architexte donné et leur adéquation pour faire transiter une information dans un niveau perceptif donné. La construction de ce modèle d'architexte est relative à une analyse sémiologique sachant que les capacités de l'architexte sont conditionnées par les dispositifs disponibles ou implantés (cf. Chapitre 10 -2.5.c page 199). Les dispositifs personnalisés ainsi que la mise en relation d'un architexte avec notre moteur de restitution documentaire sont du ressort d'un informaticien.

#### *1.2.a Le moteur de restitution documentaire*

Le moteur de restitution documentaire se compose de quatre étages indépendants (cf. Figure 91) :

- La manipulation de documents,
- Le filtrage,
- L'interaction,
- La mise en forme.

La *manipulation de documents* s'attache à associer, fusionner différents documents structurés. Chaque document structuré constitue un document de capitalisation qui va pouvoir être couplé à d'autres pour en faciliter la mise en relation.

Le *filtrage* utilise des processus de transformation génériques permettant d'opérer un élagage simple dans les documents structurés traités de façon à ne présenter que l'information utile.

L'*interaction* correspond à un étage de traitement qui va intégrer au sein des documents des mécanismes d'interaction. Ces mécanismes pourront permettre de naviguer dans le document, d'en changer dynamiquement la forme ou encore d'en modifier le contenu.

Enfin, la *mise en forme* s'appuie à la fois sur les bibliothèques génériques permettant d'accorder le document structuré avec l'architexte choisi, ainsi que sur la spécification d'un langage traduisant le passage de la sémantique du contenu vers les moyens de visualisation de l'architexte.

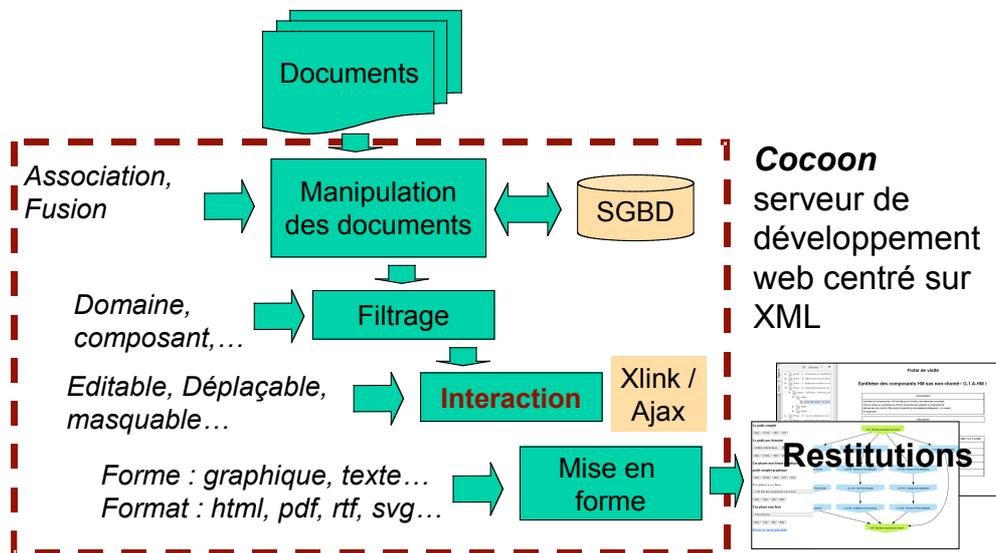


Figure 91 Architecture globale de la maquette implantée

### 1.2.b L'implémentation du moteur de restitution documentaire

La proposition, à la fois méthodologique et instrumentale, que nous avons décrite dans le chapitre précédent a fait l'objet d'une mise en œuvre s'appuyant sur le standard XML (cf. Chapitre 1 -2.2.b page 29). Ce standard permet une structuration fine de l'information via la définition de langages balisés qui pourront être logiques ou généralistes (cf. Chapitre 3 -1.3.b page 67). Il instrumente la création de documents structurés qui pourront alors être validés via des modèles de document sous la forme de DTD ou Xschema.

L'intérêt du standard XML est qu'il constitue une norme de structuration quasiment généralisée de nos jours. Il véhicule également un éventail de solutions logicielles existantes permettant de valoriser les informations répondants à ce standard. Dans notre étude, les développements sont ainsi appuyés par une solution proposée par la société *Apache* : la plateforme de développement d'applications web Cocoon (<http://cocoon.apache.org>). Ce serveur, qui constitue une solution libre, permet la mise en œuvre et l'enchaînement de traitement sur les documents structurés en XML sans engager un effort de développement trop important et de transformation XSL.

De plus, nous avons adapté Cocoon à nos besoins spécifiques puisque nous l'avons complété par d'autres bibliothèques logicielles ou composants tel que :

- Graphviz,
- Toolkit ajax.

La bibliothèque *Graphviz* (J. Ellson et al. 2003) a déjà été introduite au Chapitre 10 -2.5. page 193. Il s'agit d'une bibliothèque logicielle de calcul de visualisation de graphe. Cette bibliothèque intègre des techniques classiques de visualisation de graphe avec un temps de calcul rapide. Dans notre maquette, Graphviz est utilisé pour produire une visualisation de graphes en SVG (Scalable Vector Graphics)(Eisenberg 2003). Ce format vectoriel d'image codé en XML nous permet de rester dans un environnement cohérent facilitant ainsi la mise en relation texte et image. Une description plus étendue des possibilités de cette bibliothèque est disponible en 0 page 292.

Les toolkit ajax pour leur part doivent faciliter la mise à disposition des documents structurés sous la forme de page web accessible via un navigateur internet. L'intérêt de ces boîtes à outils est qu'elles proposent un ensemble de « widgets », comprenons des dispositifs, prédéfinis accélérant le développement d'applications web complexes. Un exemple est par exemple le « filteringTable » proposé dans le toolkit dojo (<http://dojotoolkit.org/>). Ce dispositif correspond à une table pouvant accueillir des données et intégrant nativement les fonctions de tri par colonne ou encore des possibilités de filtrage.

Id	Name	Date Added	Date Modified	
3	Carla	04/23/02	mars 01, 2004	tincidunt ut laoreet dolore magna aliquam erat volutpat.
8	Helga	04/23/02	mars 01, 2004	adipiscing elit, sed diam nonummy nibh euismod
18	Ronald	04/23/02	avr. 23, 2002	suscipit lobortis nisl ut aliquip ex ea commodo consequat.
13	Mike	04/23/02	mars 01, 2004	<b>Lorem ipsum</b> dolor sit amet...
23	Walter	04/23/02	avr. 23, 2002	consectetuer nostrud exerci tation ullamcorper
19	Stacy	11/01/03	nov. 01, 2003	<b>Lorem ipsum</b> dolor sit amet...
24	Xerxes	11/01/03	nov. 01, 2003	consectetuer suscipit lobortis nisl ut aliquip ex ea commodo consequat.
4	David	11/01/03	juin 15, 2005	Ut wisi enim ad minim veniam, quis
6	Anna	11/01/03	juin 15, 2005	tincidunt ut laoreet dolore magna aliquam erat volutpat.

Figure 92 Exemple de Filtering Table

Nous pourrions également citer d'autres dispositifs tel que la gestion de galerie photos (cf. Figure 94), l'affichage d'une « info-bulle » (cf. Figure 93) ou encore la visualisation de graphiques.

## Features

The **width** of the tooltips is adapted automatically. The tooltips may contain plain text as well as HTML, such as **images**, line **breaks**, **tables**, lists etc.

If you like to, you can **customize** these crossbrowser JavaScript Tooltips and their **behavior** in multiple ways. You can change the default configuration values inside the library itself (section GLOBAL TOOLTIP CONFIGURATION), or configure each tooltip individually by passing **commands** to the onmouseover eventhandlers.

Moreover, there are **Extensions** provide even more options to customize tooltips.

This Tooltip DHTML Library is even capable of dynamically converting arbitrary HTML elements to tooltips, which means you can put really important stuff into tooltips since HTML content is relevant to web search engines. **Documentation.**



Figure 93 Exemple « d’info-bulle » (source : [http://www.walterzorn.com/tooltip/tooltip\\_e.htm](http://www.walterzorn.com/tooltip/tooltip_e.htm))

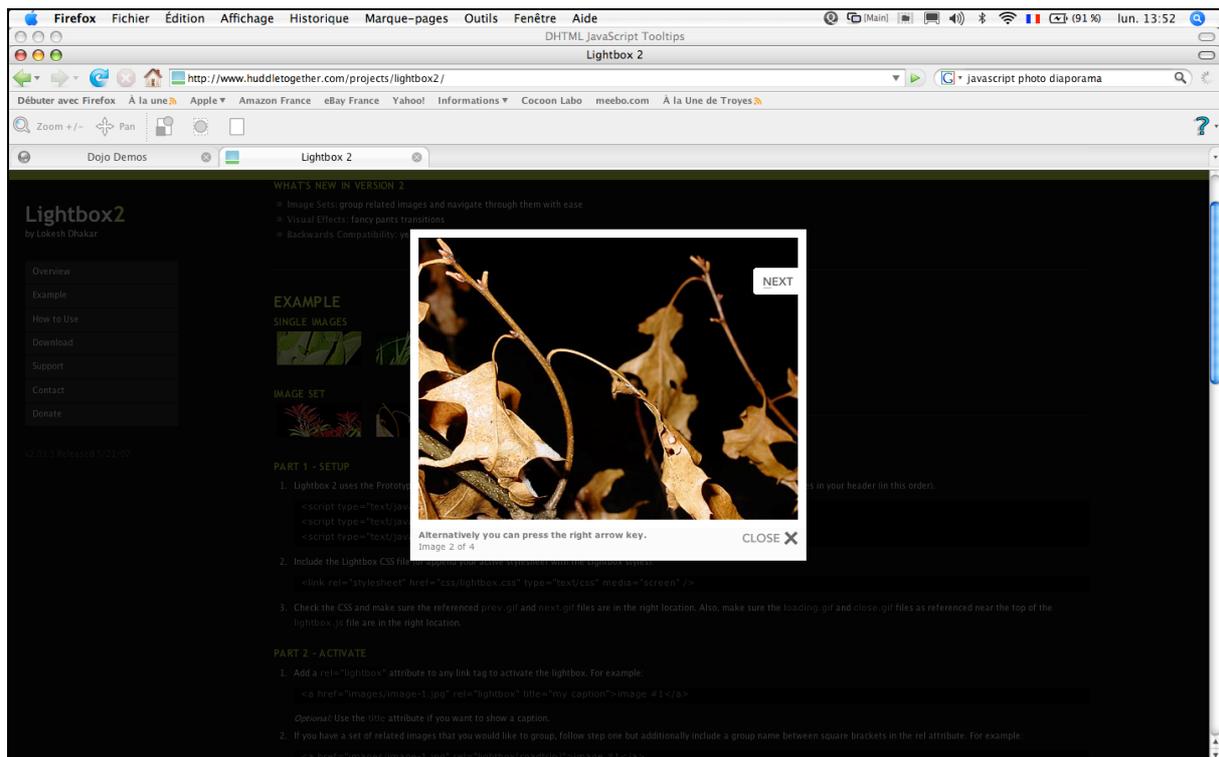


Figure 94 Exemple de galerie photos (source : <http://www.huddletogether.com/projects/lightbox2/> )

Nous avons développé un ensemble de bibliothèques génériques, constituées de feuilles de transformation XSLT, permettant d’exploiter les possibilités de Graphviz et des toolkit ajax

directement au sein de cocoon. Certaines bibliothèques ont également été développées afin de pouvoir transformer un document structuré pour qu'il devienne lisible par un ensemble d'architextes cibles tels que :

- Lecteur PDF (Adobe Acrobat Reader),
- Inkscape permettant de lire des images en SVG,
- OpenOffice Writer,
- OpenOffice Calc,
- Freemind.

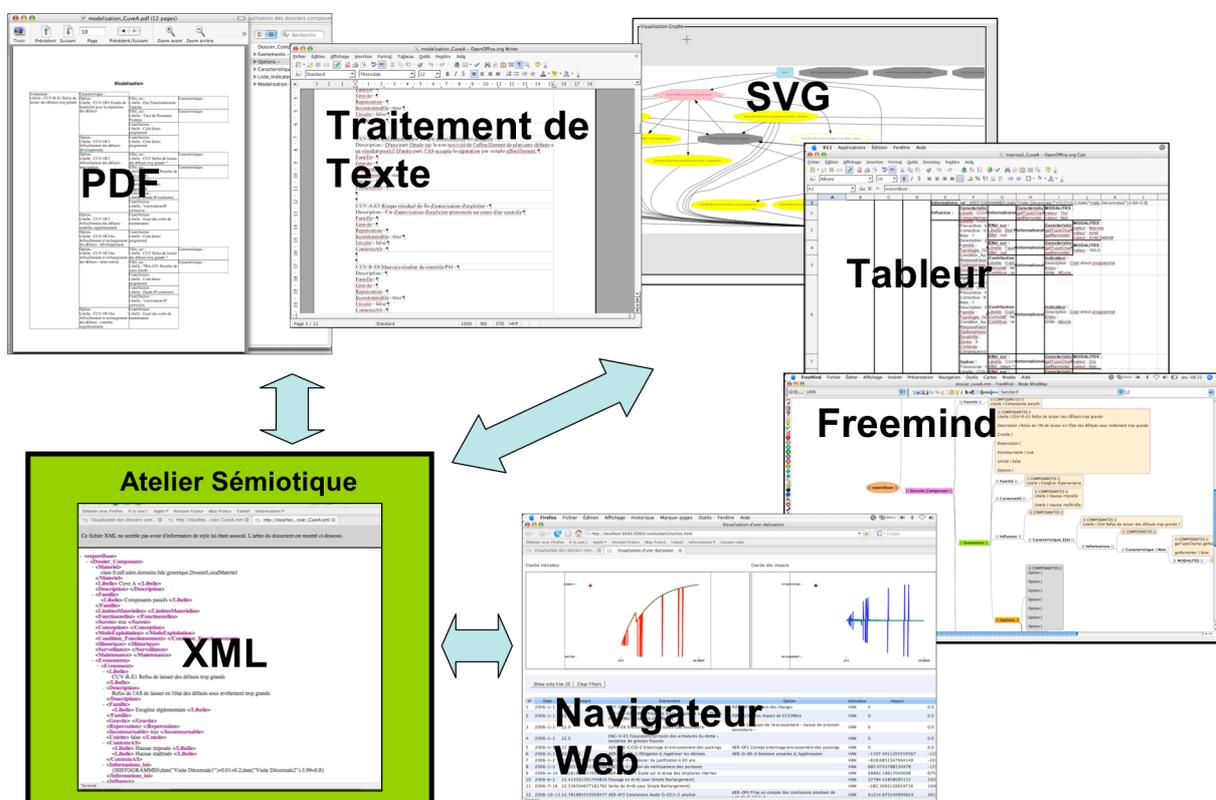


Figure 95 Un éventail de restitutions possibles

Nous présentons ici une description de la mise en œuvre de notre approche dans les deux projets réels sur lesquels s'appuie notre étude :

- Les guides de diagnostic : ce projet renvoie à une problématique de capitalisation des connaissances théoriques et des pratiques métier de diagnostic liées à de grands ouvrages hydrauliques. Ce système est construit de façon à « représenter » un support à la conservation et la diffusion d'une « représentation » des connaissances et des pratiques liées au diagnostic. (cf. section Chapitre 1 -3.2.a page 43)

- Méthodes et outils d'aide à la gestion des actifs de production : ce projet renvoie à calculer l'évolution d'indicateurs technico-économiques pour développer des politiques de maintenance à long terme. (cf. section Chapitre 1 -3.2.b page 46)

Notons bien que ces 2 projets n'ont pas uniquement représenté des cas d'application. En effet, dans notre démarche de recherche-action, ces projets ont également été des terrains d'investigation qui ont permis de définir et faire évoluer l'approche et l'instrumentation proposée dans le chapitre précédent.

## 2 - Réalisation : Les guides d'aide au diagnostic

Dans le cas d'application du guide d'aide au diagnostic, l'existant en terme d'information numérisée se limitait à une version produite via Microsoft Word puis Microsoft Visio. Dans cet exemple, nous nous sommes alors intéressés à instrumenter la tâche de capitalisation dans son intégralité. Nous proposons de porter un regard sur :

- La capitalisation,
- La restitution.

L'exemple des guides de diagnostic a été évoqué et a grandement illustré notre étude tout au long de ce mémoire. Aussi, il ne s'agit pas de reprendre l'ensemble des éléments déjà abordés. Ainsi, dans le cadre de la capitalisation, nous avons déjà évoqué l'intérêt du passage au document numérique structuré (cf. Chapitre 10 -1.2. page 172) ainsi que la manière de l'instrumenter (cf. Chapitre 10 -1.3. page 175). Nous nous concentrerons donc ici à décrire la modification du réseau d'acteurs impliqués dans la capitalisation d'une part et d'autre part les perspectives d'évolution de cette capitalisation. L'aspect restitution a également été décrit avec précision puisque l'exemple du guide a illustré notre démarche de restitution (cf. Chapitre 10 -2 - page 180). L'objet ici est donc de proposer l'éventail des restitutions produites ou en cours de développement sans nécessairement se concentrer sur le processus de construction des documents.

### 2.1. La capitalisation

Comme nous avons pu l'évoquer, initialement, la capitalisation était basée sur la rédaction d'un document. Cette rédaction résultait d'une collaboration entre experts et analyste à travers un travail de co-construction (cf. Figure 96). L'analyste s'appuyant sur des logiciels courants

avait une autonomie sur la définition de la forme de représentation de l'information produite (cf. Chapitre 8 -1 - page 148).

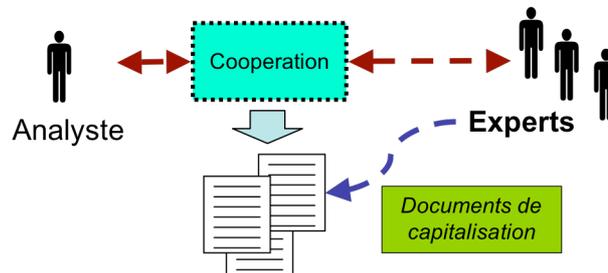


Figure 96 Une capitalisation basée sur une coopération entre experts et analyste

Les documents produits étaient alors relativement « statiques » au sens où leur mise à jour était une activité longue et manuelle. Leur production par un logiciel propriétaire rendait la pérennité du document liée à l'existence de ce logiciel comme moyen de lecture et d'écriture ou à l'existence d'une version imprimée. Cela rendait difficile le maintien et la gestion des informations qu'ils contiennent. Dans le cadre de documents techniques, qui présentent de nombreux liens et renvois entre différentes parties des documents, la gestion de la cohérence est également rendue difficile. L'approche que nous proposons trouve toute sa place dans cette problématique.

En effet, elle a offert une solution très flexible de structuration de l'information sous la forme de documents structurés (cf. Chapitre 10 -1.2.a page 173). Le document structuré identifie alors avec précision le contenu du document par le biais d'éléments structuraux qui vont délimiter le contenu. Le passage au document structuré s'est fait dans un premier temps par une traduction des documents existants vers cette forme de représentation. L'analyste, montant en expertise sur cet outil, a ensuite développé sa propre stratégie d'utilisation pour enrichir et compléter le contenu de ces documents structurés. Le document structuré s'est révélé très flexible pour produire une capitalisation sous une forme très structurée en autonomie par l'analyste. Cette forme très structurée permet d'*extraire un modèle de structuration*.

Le passage au document structuré pose toutefois une problématique au niveau de la forme puisque l'analyste n'est alors plus en autonomie face à sa réalisation (cf. Figure 97). Ceci introduit une modification dans le dispositif organisationnel initial puisque l'on assiste à une double coopération entre *informaticiens, experts et analyste*, ce dernier restant au centre de ce

dispositif en médiateur, orchestrateur de la démarche de gestion des connaissances (cf. Chapitre 6 -1 - page 130). La coopération entre l'analyste et l'informaticien doit permettre de déterminer les modalités de restitution du document structuré. On peut souligner que les propriétés du document structuré (cf. Chapitre 9 -2.1. page 163) facilitent grandement cette coopération. En effet, bien que le modèle de structuration n'ait pas été pré-établi, le document structuré constitue un espace à la fois assez souple pour une utilisation par l'analyste et assez structuré pour une valorisation par l'informaticien. La coopération entre analyste et informaticien permet ainsi de définir une palette de restitutions tout aussi bien vouées à un support papier qu'à un support numérique (cf. 2.2. page 232).

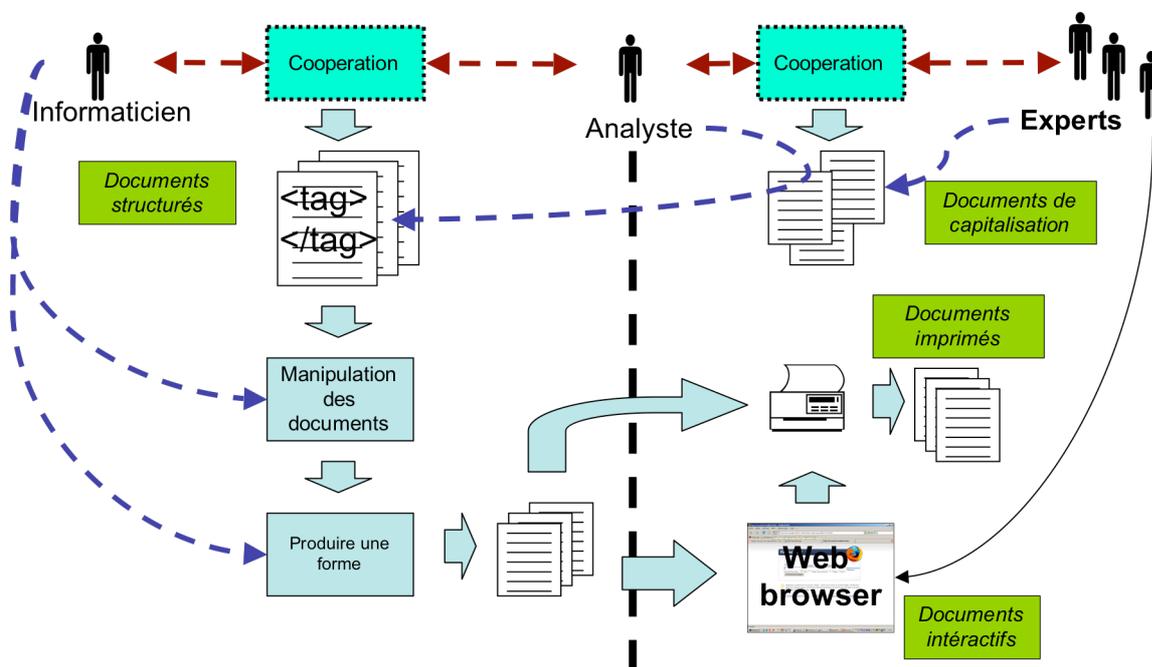


Figure 97 Une double coopération pour capitaliser et définir la forme de restitution

Le passage au document structuré ouvre d'autres perspectives d'évolution concernant le dispositif organisationnel de capitalisation. Les propriétés du document structuré permettent une indexation très fine de l'information par l'introduction d'éléments structuraux au sein du contenu. Le document est alors découpé en fragment qui peuvent faire l'objet d'annotations ou subir une modification lors de la restitution du document. Ce dernier point ouvre des perspectives très intéressantes même si nous gardons toujours à l'esprit de ne pas prendre le document structuré uniquement comme « lego » afin de limiter la dilution du document (cf. Chapitre 3 -2.2. page 76). Un exemple de personnalisation du document est d'adapter son

contenu en fonction de l'ouvrage concerné par le diagnostic (les ouvrages présentant parfois certaines spécificités, le diagnostic peut sensiblement varier).

Ainsi, lors des premiers déploiements de la maquette relative aux guides de diagnostic, les experts ont directement mis l'accent sur le fait que lors de la personnalisation par ouvrage, les photos intégrées au document produit ne correspondaient pas à l'ouvrage en question. Cette situation nous a alors amené à étudier la possibilité que le processus de personnalisation touche également cette modalité. Ainsi, nous proposons de gérer ce problème en étendant le dispositif organisationnel de capitalisation jusqu'au terrain. Notre proposition est passée par la mise en place d'un site internet de gestion de contenu (cf. Figure 98).

Ce site permet aux experts de déposer sur un espace centralisé photos et documents en rapport avec leur activité. Le dépôt passe par une indexation des éléments déposés au sens où il est préconisé d'y associer des mots-clés permettant par exemple pour une photo de préciser ce qu'elle représente et l'ouvrage où elle a été prise. L'indexation est alors sociale, à l'image des folksonomies (Ertzscheid 2006), et laissée à la discrétion des experts même s'il est possible de limiter cette indexation à un vocabulaire contrôlé. De là, il est possible de développer un mécanisme implanté permettant de personnaliser le guide à partir de ces mots-clés et des fragments délimités dans le document structuré.

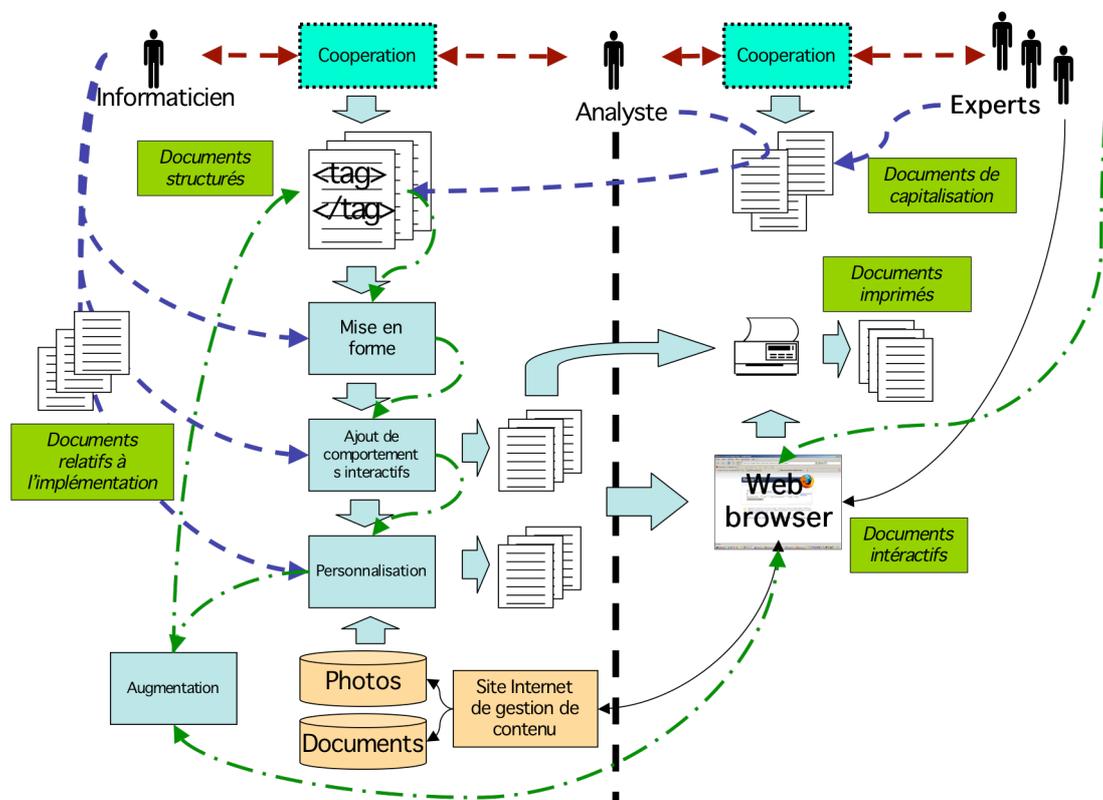


Figure 98 La capitalisation étendue aux experts via un site internet de gestion de contenu

## 2.2. La restitution

La restitution du guide joue tout d'abord un rôle primordial durant la capitalisation elle-même. Nous avons mis l'accent sur les difficultés de saisie des documents structurés (cf. Chapitre 10 -1.2.b page 174). En effet, la saisie des documents sous la forme d'un code source balisé peut se révéler être une tâche difficile en particulier lorsque le document prend de l'ampleur. Nous avons alors intégré dans la capitalisation un éditeur XML spécialisé (Morphon) permettant de transformer les balises dans une charte graphique particulière. Pour assister la saisie des documents et en particulier leur dimension structurelle, les documents peuvent également être traduits sous la forme d'une carte conceptuelle. Cette représentation graphique de type arborescente doit assister l'édition de la structure. Ce travail de restitution pour instrumenter la capitalisation a permis de développer une première bibliothèque de dispositifs relative à l'éditeur de cartes conceptuelles Freemind. Cette bibliothèque est alors indépendante de la problématique du guide et peut être utilisée sur n'importe quel document structuré.

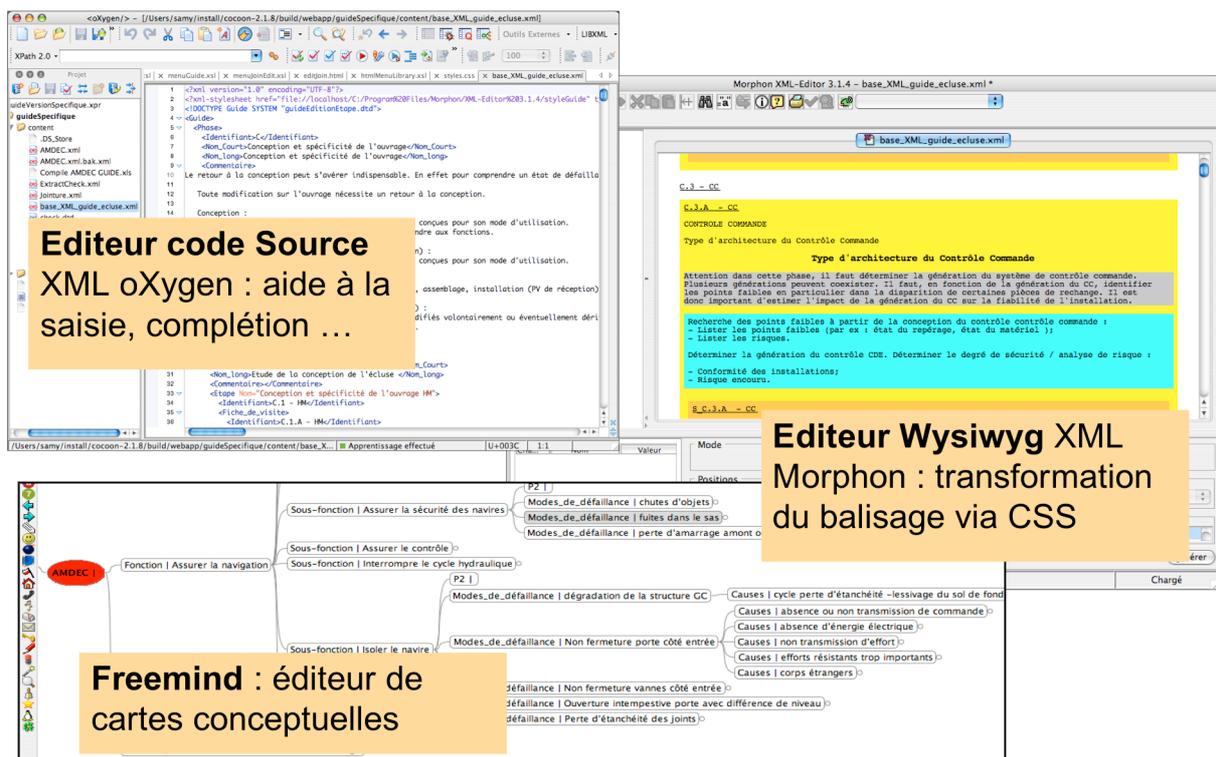


Figure 99 Restitutions pour la capitalisation

Plus largement, en s'appuyant sur un ensemble de bibliothèques permettant la transformation et l'adaptation à des architectes existants, nous proposons un éventail de modes de restitution possibles pour le guide.

### *2.2.a Multimodalité*

Nous venons de souligner dans le cadre de la capitalisation toute l'importance de pouvoir proposer un éventail de restitutions à partir d'une source unique qui se matérialise par un document structuré. C'est tout l'apport du document structuré que nous avons développé au Chapitre 9 -2.1. page 163.

Dans le cas du guide, les différentes modalités du guide que nous ciblons touchent tout aussi bien au support de manipulation du document produit qu'à la forme de l'inscription. Le support doit en effet pouvoir être papier ou numérique. En effet dans un contexte d'exploitation sur le terrain, il est encore aujourd'hui essentiel de conserver le support papier puisque les conditions d'intervention ne permettent pas toujours de pouvoir s'équiper d'un ordinateur ou généralement d'un terminal numérique adapté. De plus, en dehors d'une exploitation sur le terrain, le papier possède des qualités que l'écran ne peut égaler encore aujourd'hui. Un troisième argument est qu'une version papier peut prendre une valeur légale dans les institutions alors que le statut du support numérique n'est pour sa part pas toujours clairement défini.

#### *Multimodalité sur le plan du support*

Dans la réalisation de la version papier du guide, nous nous sommes focalisés sur le format PDF qui constitue un format de document adapté à la description de page (cf. Chapitre 3 -1.2. page 64). La technologie employée pour transformer les documents structurés en PDF est le XSL-FO (eXtensible Stylesheet Language - Formatting Objects) (Amann et Rigaux 2003). Nous avons ainsi défini une bibliothèque de dispositifs via les possibilités du XSL-FO permettant de produire titres, tableaux, d'insérer des d'images, sommaire, etc... A partir de cette bibliothèque, il est possible d'aboutir à une version du guide adaptée à une consultation sur écran mais également imprimable (cf. Figure 100).

Bien que le format PDF soit particulièrement préconisé pour l'impression sur papier dans notre approche, il constitue également un bon moyen de visualisation sur écran. Il intègre nativement des possibilités d'interaction telles que les liens hypertexte ou encore les mécanismes de signets. Nous avons ainsi la possibilité de générer dynamiquement des versions du guide PDF, qui à l'issue du calcul, aboutissent à une image statique du contenu pouvant être difficilement modifiée ou enrichie. C'est à la fois une force du format PDF puisqu'il pourra jouer un rôle de *fac simile* mais également une faiblesse puisque l'on ne

profite pas du dynamisme apporté par les possibilités du numérique. Rappelons bien ici qu'il faut manier ce dynamisme avec précautions (cf. Chapitre 3 -2.2. page 76).

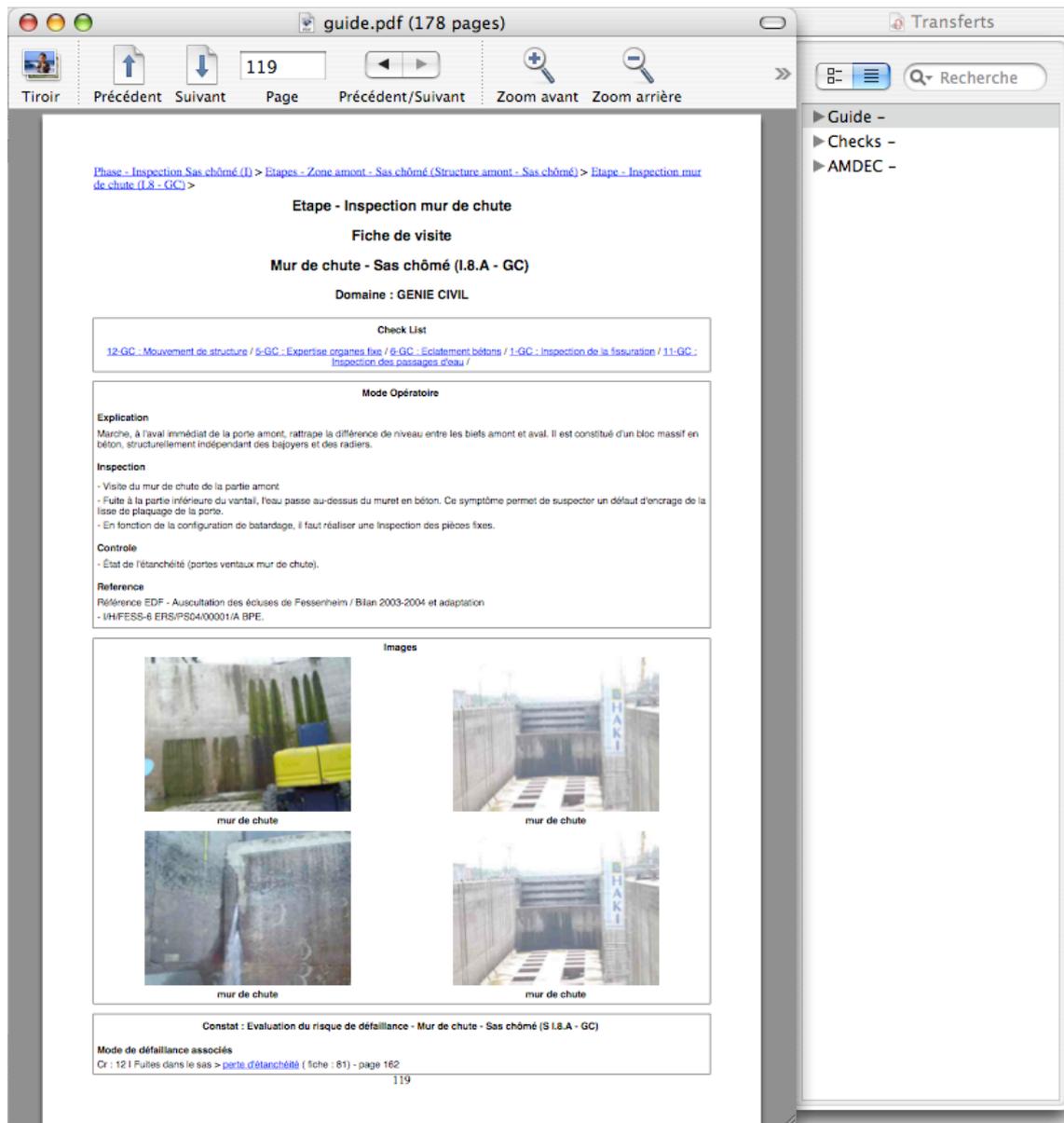


Figure 100 Une page du guide de diagnostic dans sa version PDF

Nous avons donc également développé une version du guide pouvant être consultée par le biais d'un navigateur internet (cf. Figure 103). Ce moyen offre des possibilités de déploiement réellement très étendues puisque la majorité des terminaux de consultation d'informations intègrent un navigateur de nos jours. De plus, les possibilités de visualisation d'informations sont vastes puisque qu'il est possible d'y mêler différentes modalités d'information. L'interaction est quant à elle quasiment infinie via la programmation mais peut également s'appuyer sur des « toolkits » existants (cf. Annexe 1 - page 293). Cette interaction

pourra permettre de naviguer dans le document ou d'en modifier l'aspect in vivo durant sa consultation.

### Multimodalité sur le plan de l'inscription

Une des forces du passage au document structuré est de permettre à partir d'une seule source de données de produire une multitude de formes de visualisation d'un même document. Notre approche et nos développements doivent faciliter la définition de ces différentes formes de visualisation en proposant des bibliothèques de dispositifs adaptées à la construction de document en accord avec les contraintes d'un architecte cible plus précisément d'un espace de représentation cible. La Figure 101 et Figure 102 présentent des exemples de rendu des guides de diagnostic dans le tableur et le traitement de texte de la suite bureautique libre OpenOffice.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1						<b>Fiche de visite :</b> Identifiant : D.1.A - HM Domaine : HYDRO-MECANIQUE Nom_Court : Fonctionnement hydro-méca de l'écluse Nom_Long : Fonctionnement hydro-méca de l'écluse Commentaire : L'écoute du fonctionnement sans démontage : - Placé à proximité des organes en mouvement (système de manoeuvre des portes, des vannes aq Description : Images : Check_List : Sorties :										
2			Phass : Identifiant : D Nom_Court : C Nom_Long : C Commentaire : Logigramme	<b>Etapes :</b> Identifiant : E Nom_Court : E Nom_Long : E Commentaire :		<b>Fiche de visite :</b> Identifiant : D.2.A - CC Domaine : CONTROLE COMMANDE Nom_Court : Observation du contrôle commande de l'écluse Nom_Long : Observation du contrôle commande de l'écluse Commentaire : - Impression du fonctionnement par l'exploiter - Retour des impressions de l'exploitant sur le comportement des sas, des portes, etc. Description : Images : Sorties :										
3						<b>Fiche de visite :</b> Identifiant : D.3.B - GC Domaine : GENIE CIVIL Nom_Court : Observation de la structure génie civil de l'écluse Nom_Long : Observation de la structure génie civil de l'écluse Commentaire : Elle est réalisée par l'exploitant. Elle est partielle car elle est réalisée <u>écluse</u> en eau (en exploitation). Elle consiste à faire une inspection des parties d'ouvrage a Description : Images : Sorties :										
4						<b>Fiche de visite :</b> Identifiant : E.1.A - HM Domaine : HYDRO-MECANIQUE Nom_Court : Charpente porte levante - Sas non chômé Nom_Long : Porte levante amont Charpente - Sas non chômé Commentaire : Description : Images : Check_List : Sorties :										
5						<b>Fiche de visite :</b> Identifiant : E.1.B - HM Domaine : HYDRO-MECANIQUE Nom_Court : Mécanisme de manoeuvre porte levante - Sas non chômé Nom_Long : Porte levante amont Mécanisme de manoeuvre - Sas non chômé Commentaire : Nécessite un nettoyage des organes. Prévoir des manoeuvres de porte. Cette fiche est rédigée dans le cas d'une porte amont levante manoeuvrée par câble sa Description : Images : Check_List :										

Figure 101 Le guide dans OpenOffice Calc

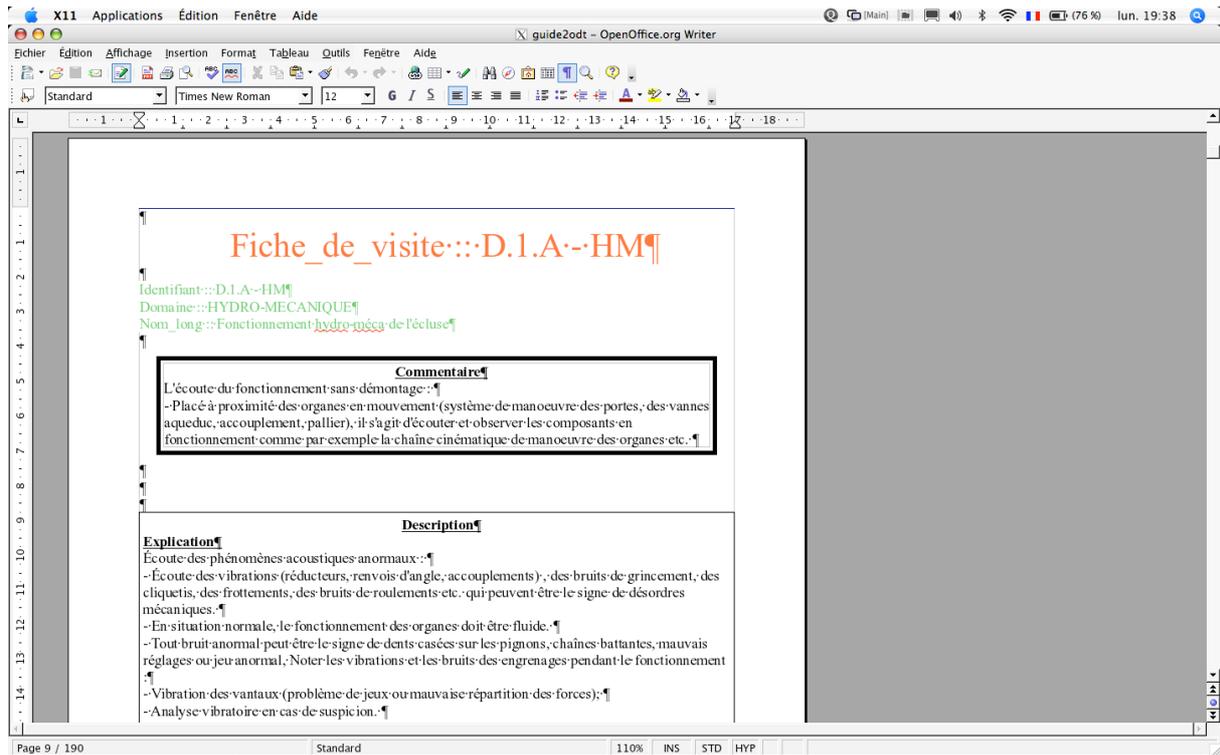


Figure 102 Le guide dans OpenOffice Writer

Dans la section précédente, nous avons évoqué certains dispositifs relatifs au format de transformation XSL-FO permettant de produire des documents PDF. Nous avons d'ailleurs illustré notre approche de restitution avec ce format au Chapitre 10 -2.5. page 193. Pour le cas du navigateur internet, nous présentons un exemple de restitution multimodale d'un document dans la Figure 103. Cette figure présente un document constitué de 2 espaces de représentation distincts qui ont été instanciés à partir d'un même document structuré.

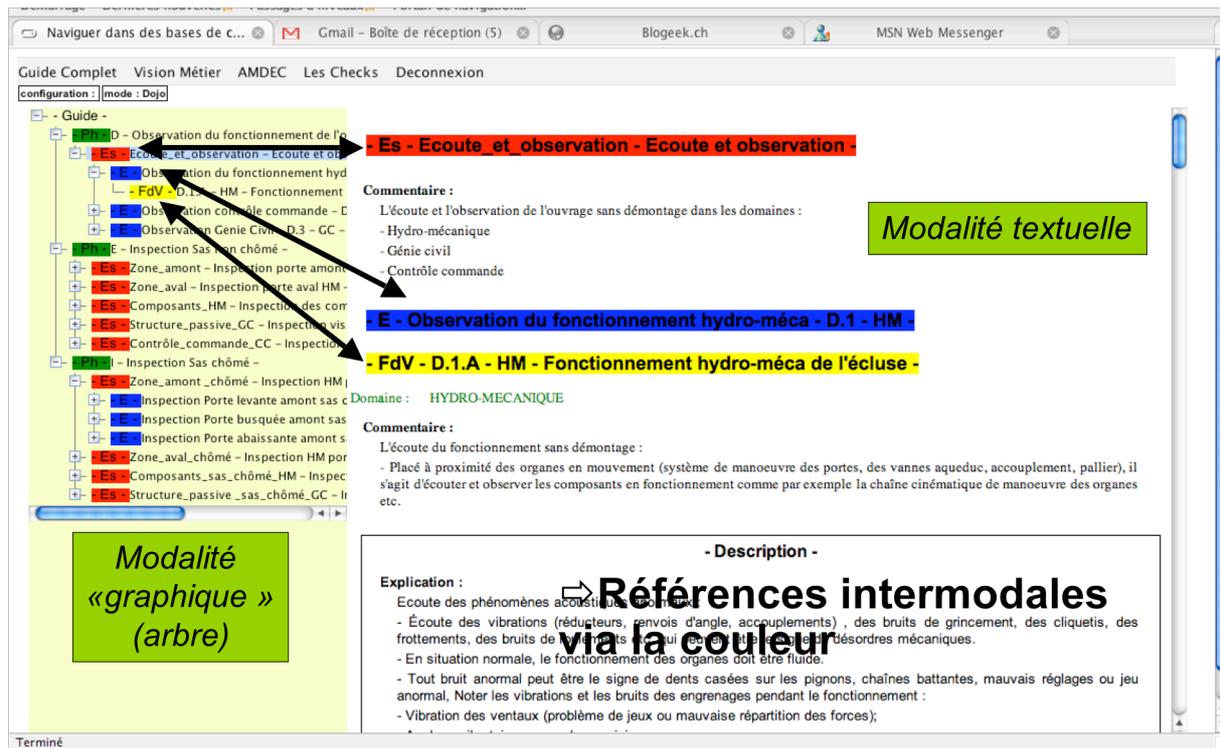


Figure 103 références intermodales

Le premier espace correspond à une arborescence de navigation qui a été construite sur la base de dispositifs en relation avec le toolkit de développement web Dojo (cf. 0 page 292). Le second espace est une représentation textuelle du contenu du guide. Le fait que ces 2 espaces soient instanciés à partir du même document facilite leur mise en relation. Ainsi, les nœuds de l'arborescence permettent de naviguer dans le document. De même, le choix des variables visuelles, dans la construction des 2 modes de visualisation, est cohérent au sens où une couleur dans un espace représente le même concept ou la même notion dans l'autre. On introduit ainsi dans la construction du document directement des références intermodales (Bestgen et Dupont 2003) qui doivent faciliter la lecture du contenu et le passage d'une modalité à une autre. Sur la Figure 103, c'est la couleur qui fait office de référence intermodale mais elle aurait tout aussi bien pu être remplacée par une forme comme une icône ou encore une étiquette numérotée.

La modalité graphique est une modalité importante puisqu'on lui associe souvent une facilité de lecture ou encore une aptitude à favoriser la compréhension. La génération de graphe est un cas particulier de représentation graphique qui nous a également permis d'illustrer notre approche de restitution (cf. au Chapitre 10 -2.5. page 193). Comme nous l'avons évoqué au 1.2.b page 224, nous avons exploité la bibliothèque de génération de graphes Graphviz afin de produire une visualisation graphique du contenu des guides de diagnostic. La Figure 104 et la

Figure 105 présentent un aperçu des résultats. Les représentations graphiques telles que nous les présentons, répondent à notre approche de restitution dans leur construction mais sous un mode supervisé (cf. Chapitre 10 -2.6. page 203).

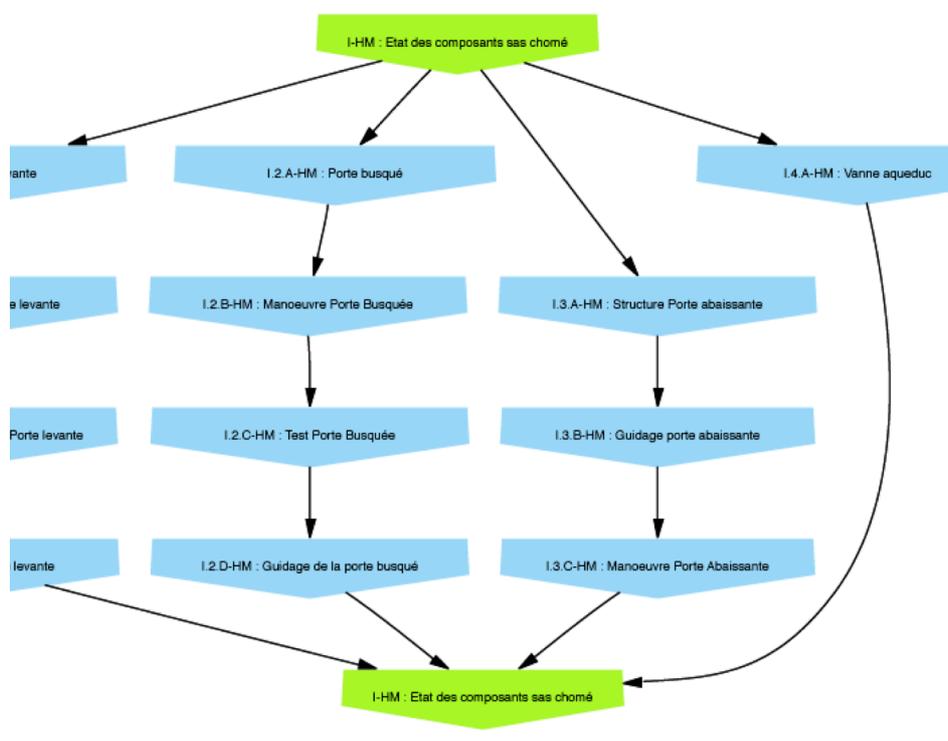


Figure 104 Exemple de diagramme

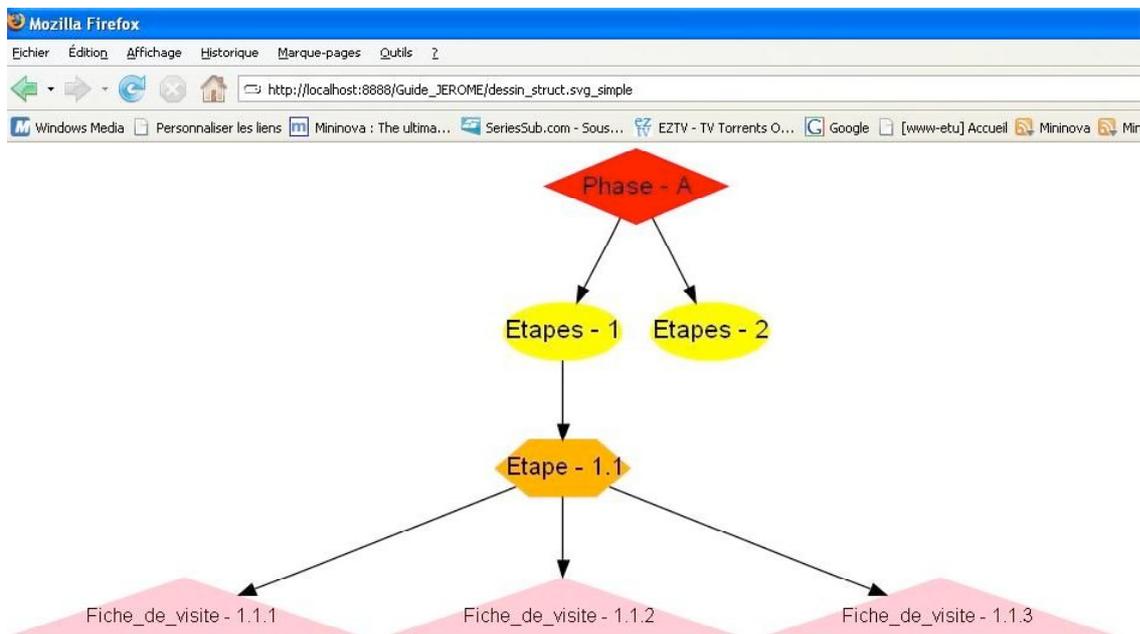


Figure 105 Visualisation sous la forme d'un diagramme accesible au sein d'un navigateur web

## 2.3. Retours d'expérience

Les guides de diagnostic se sont révélés être le principal terrain de développement et d'expérimentation de cette étude. Dans le cadre de ce projet, nous avons eu l'opportunité de développer notre approche méthodologique ainsi que de concevoir des outils opérationnels permettant d'instrumenter capitalisation et restitution des connaissances de diagnostic. Il est d'ailleurs important de souligner que la plateforme a fait l'objet d'un déploiement, la rendant accessible à un ensemble d'utilisateurs au niveau du terrain.

### 2.3.a *Le point de vue de l'analyste*

Notre approche dans ce cas d'application a permis de répondre à certaines difficultés tout en en créant de nouvelles. Dans le cadre de la capitalisation, le passage du document destiné à l'impression à une version numérique structurée a considérablement modifié l'activité de l'analyste. Si dans un premier temps, sa tâche a été de transférer une copie du guide existant dans un document structuré, il a ensuite acquis une expertise lui permettant de faire évoluer le document en autonomie, au moins sous sa forme en code source. Lors d'un entretien, il laisse d'ailleurs échapper 2 évocations importantes :

- « *La saisie dans l'éditeur xml n'est pas si différente d'un traitement de texte* »
- « *L'utilisation de xml aujourd'hui modifierait ma pratique d'entretien d'expert* »

L'évocation relative à la saisie ne doit pas être prise totalement au pied de la lettre. Nous avons d'ailleurs évoqué le problème de la saisie de document structuré au Chapitre 10 -1.2.b page 174 et la nécessité d'être en mesure de fournir un éventail de modes de saisie (cf. Chapitre 10 -1.3. page 175 et au 1.1. page 214). Ceci étant notre observation montre qu'en début de capitalisation un analyste non informaticien a pu s'adapter à la saisie en code source de documents structurés et y développer une stratégie d'utilisation adaptée à son activité.

Cette montée en expertise de l'analyste non informaticien nous amène à l'autre évocation de l'analyste. A mesure de l'avancée dans le projet, l'analyste a acquis une meilleure vision de ce qu'il était possible de réaliser en terme de traitement de contenu ou de mise en forme par exemple. Ceci n'est pas uniquement le fait de la montée en compétence de l'analyste mais également en relation avec la coopération avec un informaticien dans la définition des modes de restitution. Cette meilleure compréhension de l'analyste l'a amené à évoquer qu'il modifierait sa méthode d'entretien certainement au regard de l'identification des informations mais pourquoi pas en introduisant durant l'entretien le document structuré, pas

nécessairement sous une forme en code source, mais comme support de coopération *in vivo*. Par exemple, lors de certaines réunions réunissant experts et analyste, ce dernier opérait des ajouts ou des modifications de contenu dans le guide durant la réunion, ce qui permettait de générer une version du guide directement et ainsi de présenter le résultat en situation.

La principale difficulté reste toutefois la forme visuelle que vont prendre les documents structurés puisque si un analyste a pu s'accommoder dans une certaine mesure à une visualisation en code source, il n'en sera pas forcément de même pour des experts. Des solutions d'édition telles que l'éditeur de cartes conceptuelles Freemind ou plus simplement le tableur Excel, semblent dans ce contexte plus adaptées pour l'ensemble de nos acteurs. Dans le cas du tableur, il s'agira toutefois qu'ils saisissent leurs informations conformément au mode d'écriture que nous avons identifié (cf. Chapitre 8 -1.2. page 151). Ceci ne devrait pas constituer un obstacle du fait que nous avons constaté qu'il s'agissait d'une pratique d'écriture en fait récurrente.

### *2.3.b Le point de vue de l'informaticien*

Dans ce cas d'application, il apparaît que la pratique de l'informaticien est sortie des sentiers battus. En effet, nous avons préféré à une approche classique qui tend à anticiper la structuration des données, une approche plus ancrée dans le paradigme *extreme programming* (Casabianca 2005). Ainsi, dans la mise en œuvre de nos documents structurés nous n'avons pas souhaité définir a priori un modèle de document préférant que celui-ci émerge de l'écriture d'un document bien formé rédigé en autonomie par l'analyste.

Bien évidemment cette manière de procéder n'est pas sans poser de difficultés puisque une écriture que nous qualifierons de constructiviste tend à produire une information redondante pouvant introduire des incohérences ou encore être difficile à mettre à jour. Il s'agissait de difficultés auxquelles on était déjà confronté dans la rédaction d'un document classique voué à l'impression dès lors que celui-ci prenait de l'ampleur. Dans ce cas, il s'agissait alors d'opérer une relecture manuelle pouvant se révéler fastidieuse et complexe.

Notre expérience nous a permis de voir que le document structuré constitue un support d'inscription accessible à un non informaticien à la différence d'une base de donnée qui nécessite une certaine expertise ne serait que pour créer les tables de données. Le document structuré représente également un support et une forme d'inscription très structurés et aisément manipulables par l'informaticien. Ainsi, en s'appuyant sur une collaboration entre

analyste et informaticien, il est possible de converger vers une structure de donnée ancrée dans le document se révélant peu redondante et relativement cohérente. Ceci passe bien évidemment par des opérations de transformation du document qui vont par exemple factoriser un certain nombre de briques élémentaires dès lors que l'on identifie qu'elles sont répétées de nombreuses fois.

Plus généralement, le document structuré, en s'appuyant sur le standard XML, constitue une solution souple de manipulation et de restitution de données. Les documents peuvent ainsi passer d'une forme de visualisation à une autre, d'un support de diffusion à un autre ce qui permet de proposer un éventail de solutions sans engager un coût de développement humain et financier trop important. Dans le contexte où nous évoluons, ceci offre la possibilité de proposer de manière simple des maquettes permettant de mieux présenter ce qui est réalisable ou non.

### *2.3.c Le point de vue du « terrain » : experts et commanditaires*

Le passage d'un document papier, ou plus exactement d'un document numérique voué à l'impression, à un document numérique structuré a permis d'élargir les possibilités du guide. Tout d'abord, la souplesse de restitution du document que nous avons évoquée précédemment permet de proposer des solutions concrètes très rapidement aux experts et commanditaires qui constituent les clients au regard de la division Recherche et Développement. Dès lors, durant les réunions de l'ensemble des acteurs de la démarche de gestion des connaissances engagée, on ne débat pas sur des concepts mais bien sur une solution tangible qu'il faut adapter et faire évoluer pour qu'elle réponde aux besoins. Ainsi, plutôt que de chercher à définir un cahier des charges a priori sur une potentielle solution adaptée, il est possible de construire la solution au fur et à mesure voir même d'imaginer la créer pour partie *in vivo* avec experts commanditaires et utilisateurs.

Au regard de la solution produite, plusieurs avantages peuvent être soulignés. Tout d'abord, le caractère centralisé de la solution. Les documents structurés produits constituent une réelle base de données du savoir expert articulant à la fois domaines de compétences et points de vue. La mise en relation des différents domaines constitue un élément notable puisqu'on est le plus souvent confronté à une documentation plutôt spécifique à un domaine. La mise en relation des points de vue permet d'accéder au contenu selon une vision métier ou une vision plus fonctionnel tout en conservant la possibilité de passer de l'un à l'autre. Ainsi, le guide produit, bien que sa rédaction ne soit pas d'un niveau de précision très important, constitue

une réelle solution d'organisation de connaissances telle que nous avons pu l'évoquer au Chapitre 1 -2.3.c page 35. Les experts lors de nos derniers entretiens ont ainsi pu évoquer le potentiel du guide qui, du fait de sa conception, permet de lier de nombreuses ressources existantes qu'ils consultent dans leur activité. Il peut s'agir de documents, de résultats de calculs, de nombreux médias tels que des photos ou des vidéos. De là, le guide pourrait devenir un point d'accès à de nombreuses ressources existantes en se positionnant comme une solution à la désorientation dans des espaces numériques de plus en plus grands tels que peut l'évoquer Bachimont (cf. Chapitre 1 -2.1. page 26).

Dans une logique plus opérationnelle, le guide offre également un réel potentiel pour instrumenter les activités métier. Le commanditaire a ainsi soulevé l'intérêt de valoriser les guides de diagnostic en les associant à une Gestion de Maintenance Assistée par Ordinateur (GMAO). En effet, ce type de système qui en premier lieu offre une vision de l'ouvrage essentiellement fonctionnelle basée sur la mise en place de capteur pourra profiter du contenu du guide qui bien que certainement plus factuel reflète une expertise plus centrée sur le savoir expert pouvant ainsi appuyer les décisions.

Un autre aspect plus opérationnel s'est dégagé à la suite d'une incohérence liée à la mise en place de photos dans le guide. Le problème était que lors de la génération du guide dans une version spécifique à un ouvrage particulier, les photos introduites dans le contenu ne correspondaient pas nécessairement à l'ouvrage en question. Il s'agissait alors de mettre en place un mécanisme permettant de sélectionner une photo adaptée. Nous avons alors proposé une solution (en cours de développement) permettant aux experts de déposer photos et documents relatifs à un ouvrage particulier sur un site de gestion de contenu (cf. 2.1. page 228). A partir de cette instrumentation, il est possible d'imaginer que le guide de diagnostic puisse intégrer lors de sa génération l'ensemble des photos correspondant à un ouvrage particulier permettant ainsi d'illustrer l'évolution des parties de l'ouvrage en question.

C'est tout l'intérêt du passage au document numérique structuré puisque le guide de diagnostic ne correspond plus à un document figé essentiellement voué à l'impression. Il s'agit d'une ressource « vivante » pouvant évoluer. Il peut être enrichi sans un effort de réédition trop important. Il peut être adapté selon les spécificités d'un ouvrage particulier ou selon une activité métier particulière. Nous avons ainsi évoqué la possibilité de réorganiser le contenu par rapport à la tâche d'un technicien effectuant une observation régulière de l'ouvrage. L'idée serait d'offrir une interface permettant de définir comment on va

diagnostiquer l'ouvrage, en particulier l'ordre des composants observés, pour produire un document adapté, c'est-à-dire filtré et réordonné, à l'action de maintenance en question.

### 3 - Méthodes et outils d'aide à la gestion des actifs de production

Dans le projet relatif aux Méthodes et Outils d'aide à la gestion des actifs de production, il s'agissait de s'intégrer dans un environnement déjà existant de capitalisation et de restitution de connaissances. Dans ce rapport, nous avons souligné les limites et difficultés posées par la solution existante (cf. Chapitre 1 -3.2.b page 46 et Partie III - page 127). Notre positionnement n'a pas été de remettre en cause l'existant mais plutôt de travailler en complémentarité et prolongement de cette solution.

L'idée a été de proposer le document structuré comme une interface entre les utilisateurs et le système à base de connaissance pour instrumenter les activités de :

- Lecture du contenu de la base de connaissances,
- Lecture des résultats produit par le système,
- Saisie dans la base de connaissances.

Dans cette perspective, nous avons développé un ensemble de modèles de documents structurés en accord avec les structures de données mises en jeu dans le système à base de connaissances auquel nous étions confrontés. Ces modèles de documents représentent alors un format d'échange standard permettant l'exportation et l'importation d'informations de la base de connaissances sous la forme de documents structurés. Ces documents structurés peuvent alors profiter de notre plateforme pour être adaptés et rendus lisibles par un traitement de texte, un lecteur de fichier PDF, un tableur, un éditeur de cartes conceptuelles ou encore un navigateur web (cf. Figure 106).

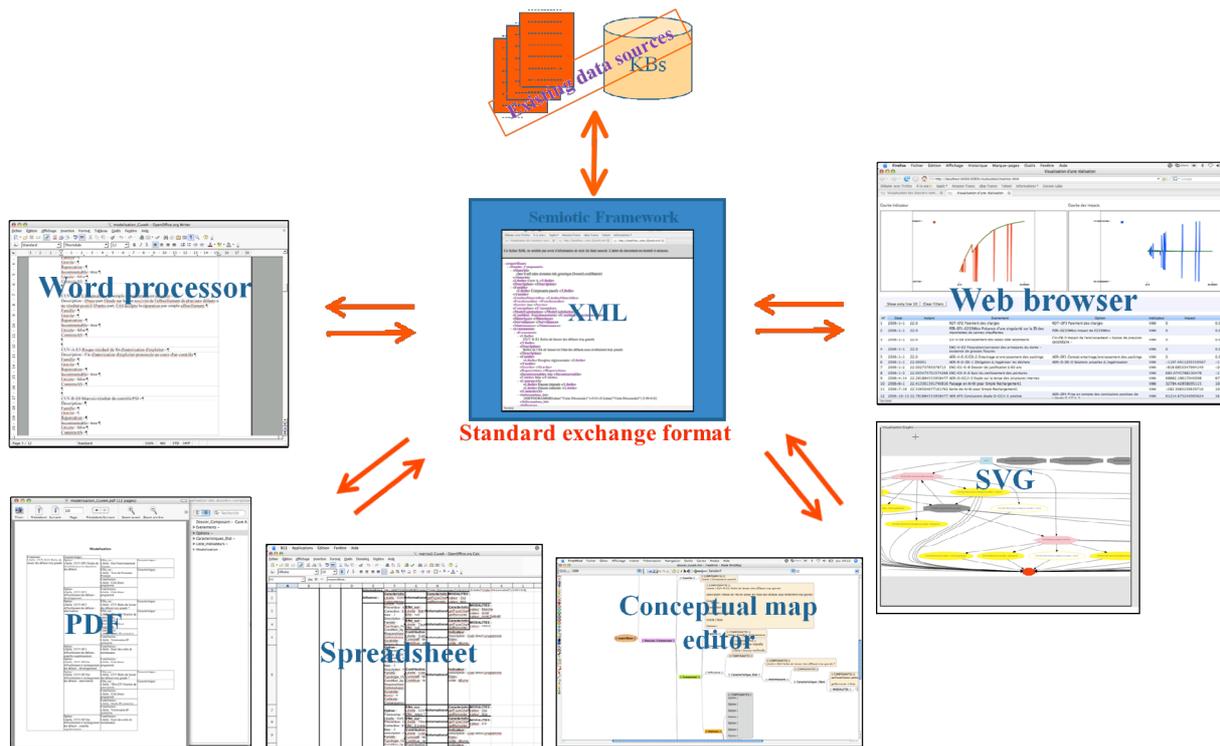


Figure 106 Le document structuré constitue un pivot entre un système existant et un éventail d'architectes

Dans les deux sections suivantes, nous décrivons la mise en œuvre de notre approche dans la capitalisation et la restitution.

### 3.1. En capitalisation

Dans le cadre de ce projet, la méthode de capitalisation était pré existante. Elle est basée sur la rédaction de documents de capitalisation dont la structure et l'organisation du contenu sont très cadrées. La capitalisation intègre également une étape de saisie au niveau d'un système informatique dans un langage formel. Toujours dans l'idée de travailler en complémentarité de l'existant, nous avons défini un modèle de document structuré intégrant les contraintes du langage de modélisation formel.

#### 3.1.a Le document structuré comme espace de saisie formel

La définition d'un modèle de document associé à un éditeur XML adapté permet de cadrer l'écriture de documents structurés qui vont pouvoir ensuite être exporté vers la base de connaissances originale. Le document structuré se pose alors en alternative des interfaces existantes, ces dernières se matérialisant majoritairement sous forme de formulaires. Nous avons introduit les limites des interfaces de type formulaire au Chapitre 8 -2.1. page 154. Le



d'éléments structuraux à remplir (cf. Figure 110). Le contenu a été complété par des éléments devant obligatoirement figurer dans le document structuré, tel que précisé dans le modèle de document, dès lors que l'on insère un nouvel événement.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE exportBase SYSTEM "file:/Users/samy/install/cocoon-2.1.9/build/webapp/EDEN/content/dossiers/dossierSaisie.dtd">
<exportBase>
  <Dossier_Composant>
    <Materiel></Materiel>
    <Libelle></Libelle>
    <Description></Description>
    <Famille></Famille>
    <LimitesMaterielles></LimitesMaterielles>
    <Fonctionnelles></Fonctionnelles>
    <Surete></Surete>
    <Conception></Conception>
    <ModeExploitation></ModeExploitation>
    <Condition_Fonctionnement></Condition_Fonctionnement>
    <Historique></Historique>
    <Surveillance></Surveillance>
    <Maintenance></Maintenance>
    <Evenements>
      |
    </Evenements>
  </Dossier_Composant>
  <Caracteristiques_Etat>
    <Caracteristique_Etat>
```

Figure 108 Document structuré en cours de saisie

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE exportBase SYSTEM "file:/Users/samy/install/cocoon-2.1.9/build/webapp/EDEN/content/dossiers/dossierSaisie.dtd">
<exportBase>
  <Dossier_Composant>
    <Materiel></Materiel>
    <Libelle></Libelle>
    <Description></Description>
    <Famille></Famille>
    <LimitesMaterielles></LimitesMaterielles>
    <Fonctionnelles></Fonctionnelles>
    <Surete></Surete>
    <Conception></Conception>
    <ModeExploitation></ModeExploitation>
    <Condition_Fonctionnement></Condition_Fonctionnement>
    <Historique></Historique>
    <Surveillance></Surveillance>
    <Maintenance></Maintenance>
    <Evenements>
      |
      <Evenement
    </Evenement>
  </Dossier_Composant>
  <Caracteristiques_Etat>
    <Caracteristique_Etat>
      <Libelle></Libelle>
      <Informations></Informations>
      <Comportement></Comportement>
      <Classe></Classe>
      <Famille></Famille>
```

Figure 109 Illustration du mécanisme de complétion et de validation en contexte

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE exportBase SYSTEM "file:/Users/samy/install/cocoon-2.1.9/build/webapp/EDEN/content/dossiers/dossierSaisie.dtd">
<exportBase>
  <Dossier_Composant>
    <Materiel></Materiel>
    <Libelle></Libelle>
    <Description></Description>
    <Famille></Famille>
    <LimitesMaterielles></LimitesMaterielles>
    <Fonctionnelles></Fonctionnelles>
    <Surete></Surete>
    <Conception></Conception>
    <ModeExploitation></ModeExploitation>
    <Condition_Fonctionnement></Condition_Fonctionnement>
    <Historique></Historique>
    <Surveillance></Surveillance>
    <Maintenance></Maintenance>
    <Evenements>
      <Evenement>
        <Libelle></Libelle>
        <Description></Description>
        <Famille></Famille>
        <Gravite></Gravite>
        <Repercution></Repercution>
        <Incontournable></Incontournable>
        <Unicite></Unicite>
        <ContexteAS>
          <Libelle></Libelle>
        </ContexteAS>
        <Informations_loi></Informations_loi>
        <Influence></Influence>
        <Options></Options>
      </Evenement>
    </Evenements>
  </Dossier_Composant>
  <Caracteristiques_Etat>
    <Caracteristique_Etat>
      <Libelle></Libelle>
      <Informations></Informations>
    </Caracteristique_Etat>
  </Caracteristiques_Etat>
</exportBase>

```

Figure 110 Document structuré complété d'un ensemble de balise à remplir

Notons que nous avons illustré la saisie via un éditeur proposant une visualisation en code source. Les balises, correspondant aux éléments structuraux, pourraient tout aussi bien être traduites dans une sémantique visuelle telle que nous l'avons évoqué au Chapitre 10 -1.3. page 175 à l'image du résultat présenté sur la Figure 111.

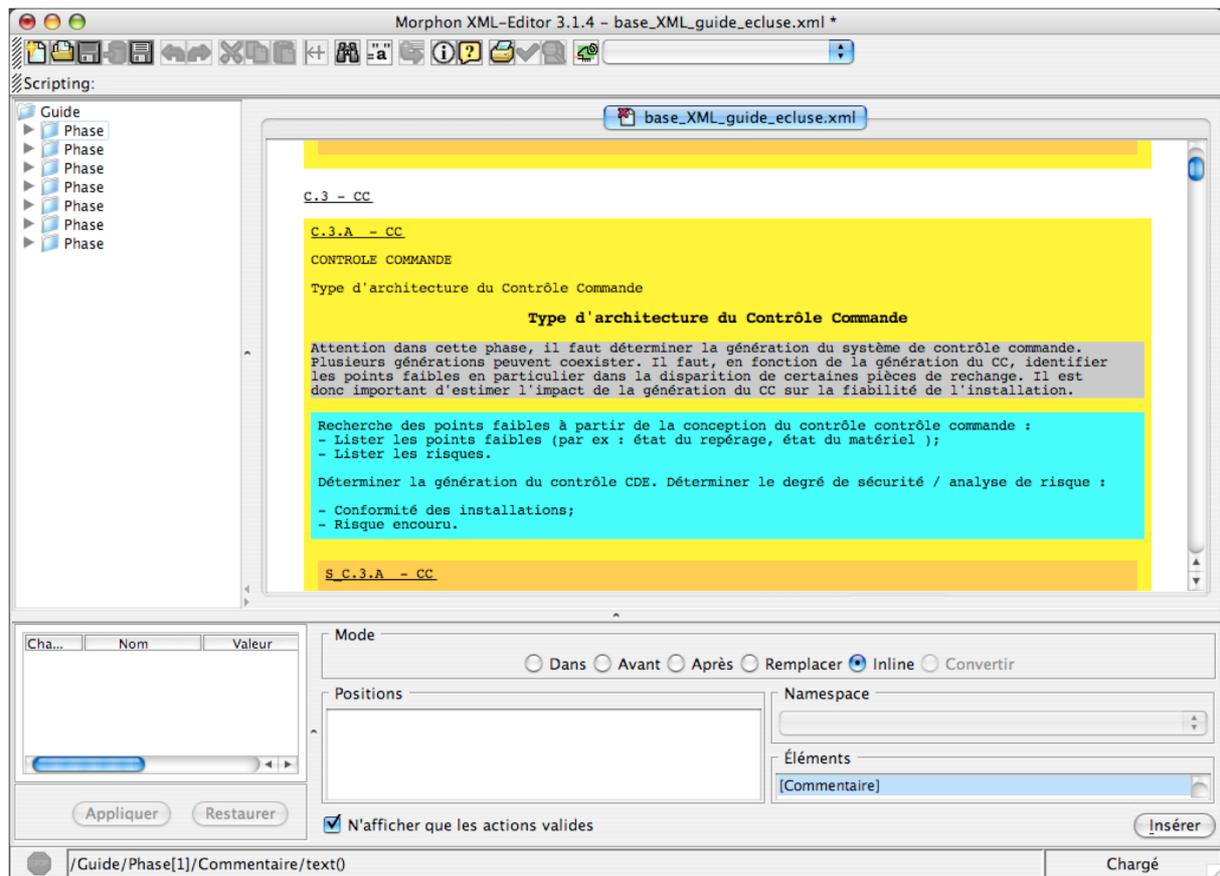


Figure 111 Traduction du balisage sous la forme de cadres colorés dans l'éditeur Morphon

Dans ce projet, nous avons ainsi prototypé d'autres modes de visualisation et de saisie basés sur des architectes existants. Nous traduisons ainsi les documents structurés au travers des possibilités des architectes via notre approche de restitution.

### 3.1.b Edition via d'autres architectes

En s'appuyant sur notre approche de restitution ainsi que les bibliothèques développées, nous proposons 3 moyens de saisie mais également de visualisation et diffusion de l'information complémentaires du système existant :

- L'éditeur de cartes conceptuelles Freemind
- Le tableur de OpenOffice Calc
- Le traitement de texte OpenOffice Writer.

L'éditeur de cartes conceptuelles Freemind offre une vision graphique, globale et interactive. Le document structuré est alors traduit sous la forme d'une arborescence. Les balises sont alors traduites sous la forme de nœuds au niveau de la carte, dont le contenu est préfixé par le nom de la balise (« || nom balise || ») que l'on colore. Cet éditeur est particulièrement adapté à

une saisie d'informations arborescente puisque le mode de visualisation qu'il propose fait bien transparaître la dimension structurelle (cf. Figure 112).

Le *tableur OpenOffice* a été mis en œuvre à l'image du mode d'écriture décrit au Chapitre 8 - 1.2. page 151. Le tableur est exploité comme un outil de modélisation et structuration d'information, au détriment de ses capacités de calcul. Il s'agit de proposer une écriture matricielle du document structuré dans un éditeur courant. Nous avons opté ici pour OpenOffice pour des raisons de facilités de développement mais l'utilisation d'une bibliothèque spécialisée aurait permis de s'interfacer avec Microsoft Excel. La Figure 113 illustre le résultat produit. Les cellules vont ainsi traduire les zones délimitées par les balises alors que les colonnes vont spécifier le niveau de profondeur auquel on se situe dans le document.

Le *traitement de texte OpenOffice Writer* doit permettre de se rapprocher du document tel qu'on peut l'appréhender couramment. Nous avons ainsi produit un document intégrant l'ensemble du contenu de la base de connaissances prenant la forme d'un inventaire des données par type. Le document produit intègre également des tableaux à l'image de ceux générés dans le tableur qui vont traduire la mise en relation d'information. Le document a été construit conformément à la trame des documents de capitalisation spécifiée dans la méthodologie de capitalisation inhérente au projet (cf. Figure 114). Il est alors très structuré conformément à ce que nous avons évoqué dans l'analyse des documents de capitalisation (cf. Chapitre 8 -1.1. page 148). Cette démarche est particulièrement intéressante puisque si dans le cas présent nous avons généré des documents à partir d'une base de connaissances déjà existante, ils pourraient devenir les moyens de saisie vers cette base de connaissances. Dans le cas qui nous occupe ici, ce mode fonctionnement pourrait permettre d'éviter une double saisie d'un nombre important d'informations qui sont à la fois inscrites de manière très structurée dans les documents puis re-saisies dans la base de connaissances. Au-delà de la double saisie, la méthode qui prendrait comme point de départ un document comme moyen de saisie, présentant une structuration identifiable par un traitement informatique afin de pouvoir en extraire les informations structurées, permettrait de réduire le fossé entre les documents et les bases de connaissances formelles.

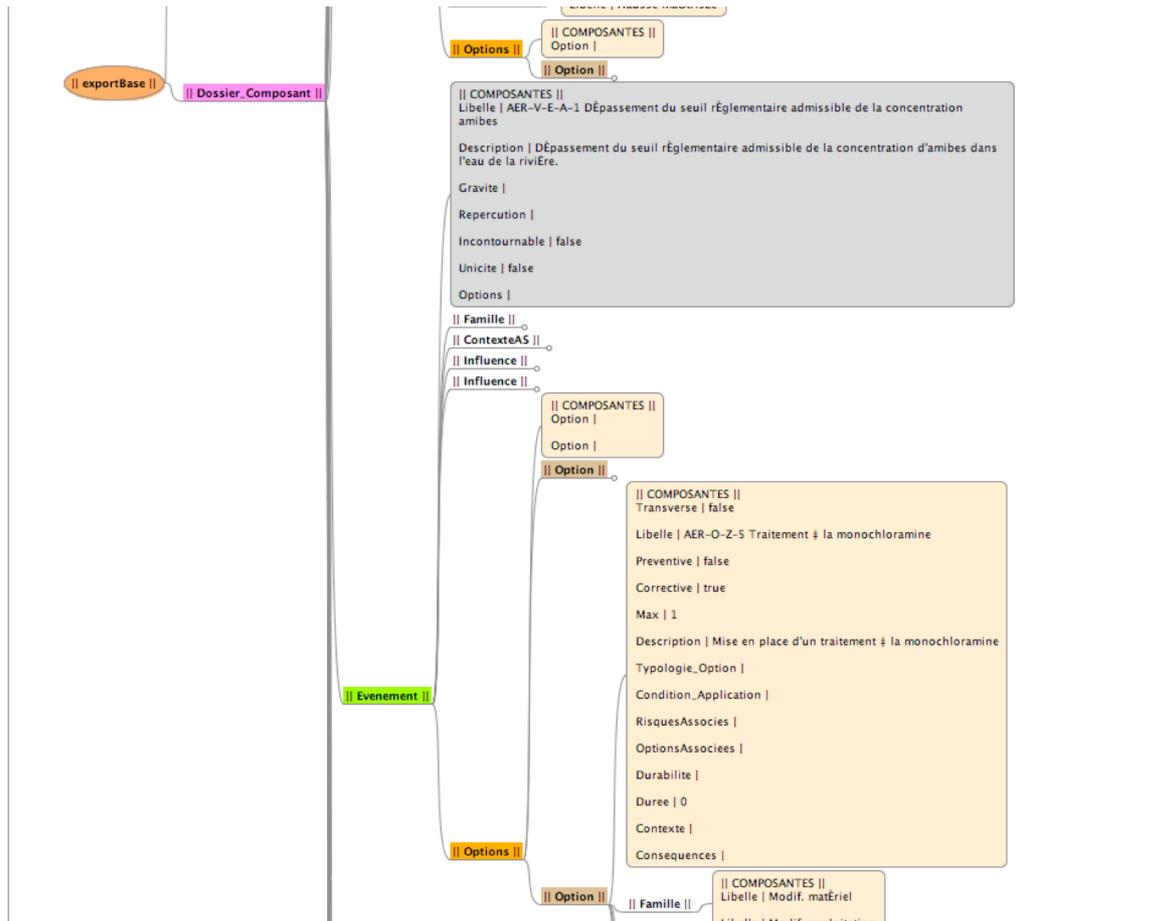


Figure 112 une carte conceptuelle produite via Freemind

	A	B	C	D	E	F	G	H	I	J	K	L	M
9						Max : 1 Description : f Famille : f Typologie_Op Condition_Ap	<b>Contribution</b> Libelle : C <sup>o</sup> ut Cumulatif : fal Contribute : re	<b>Informations</b>	<b>Indicateur :</b> Description : C <sup>o</sup> ut direct programmé Enjeu : Unite : kEuros				
10						RisquesAssoc OptionsAssoc Durabilite : Duree : 0 Contexte : Consequence	Libelle : Duré Cumulatif : fal Contribute : if	<b>Informations</b>	<b>Indicateur :</b> Description : Duré indisponibilité programmée corrective Enjeu : Unite : jours				
11						Option : Transverse : f Libelle : CUV Preventive : f Corrective : tr Max : 1 Description : (	<b>Contribution</b> Libelle : Valo Cumulatif : fal Contribute : if	<b>Informations</b>	<b>Indicateur :</b> Description : Valorisation de l'indisponibilité programmée corrective Enjeu : Unite : kEuros				
12				<b>Evenement :</b> Libelle : CUV Description : f Famille : Gravite : Repercution : Incontournabl Unicite : raise ContexteAS :	<b>Options :</b>	Famille : (	<b>Contribution</b> Libelle : Ecart Cumulatif : fal Contribute : re	<b>Informations</b>	<b>Indicateur :</b> Description : Ecart des couts de maintenance Enjeu : Unite : kEuros				

Figure 113 Capitalisation hiérarchique de l'information au sein du tableur de la suite bureautique OpenOffice

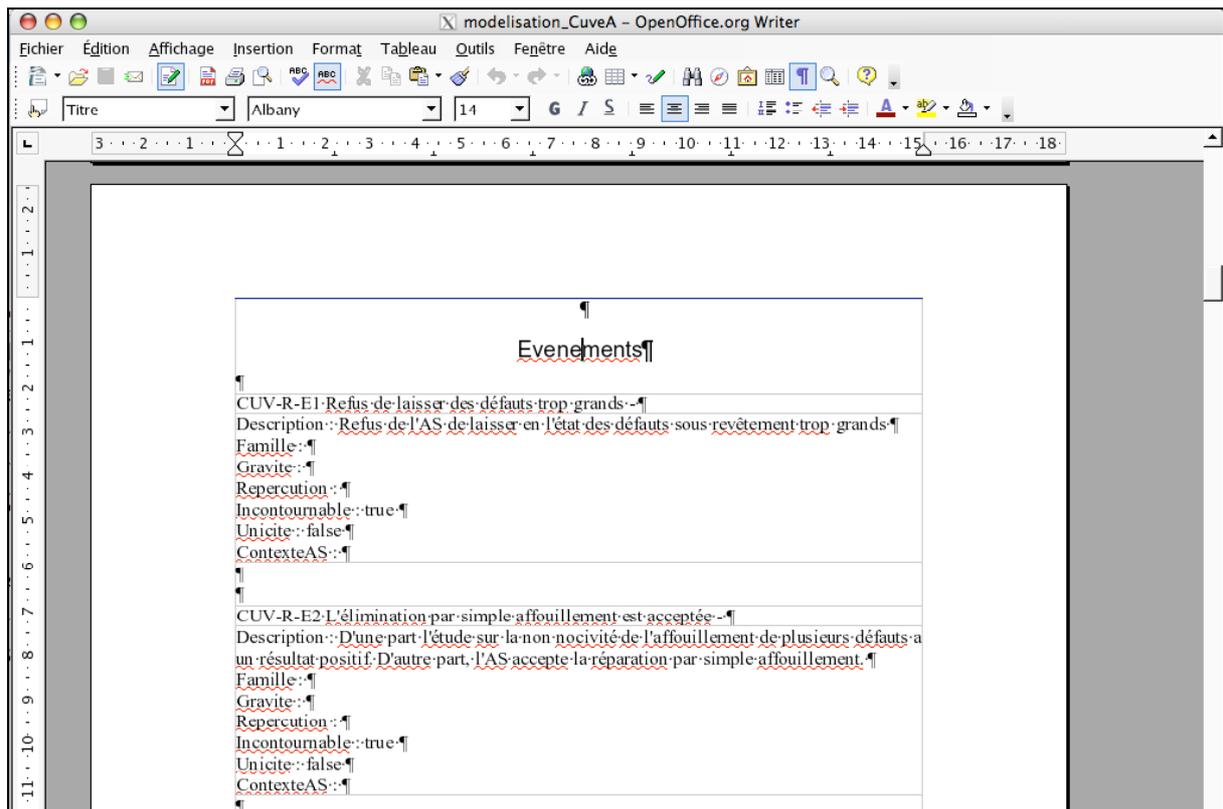


Figure 114 Document de capitalisation accessible via le traitement de texte de OpenOffice

L'utilisation de ces architextes pose toutefois une difficulté puisque contrairement à l'éditeur XML ils ne garantissent plus la validité du document produit. En effet, dès lors que l'on est dans un architexte particulier, il est a priori possible de saisir n'importe quelle balise – sous la forme définie – à n'importe quelle position : l'auteur ne profite plus d'un accompagnement dans la structuration. Ceci ne constitue toutefois pas un obstacle. Il est en effet possible de vérifier la cohérence avec un modèle formel a posteriori, voir même lors de cette vérification de cohérence faire apparaître les éléments problématiques afin qu'ils soient repris dans l'architexte en question.

### 3.2. En restitution

Dans le cadre de la restitution, nous avons développé une couche de présentation d'informations venant en complément du système à base de connaissances existant. Cette couche de présentation intègre un moteur de restitution documentaire qui va produire des documents constituant l'interface entre le système et les utilisateurs. Ce moteur de restitution s'appuie sur l'architecture technique présentée au début de ce chapitre.

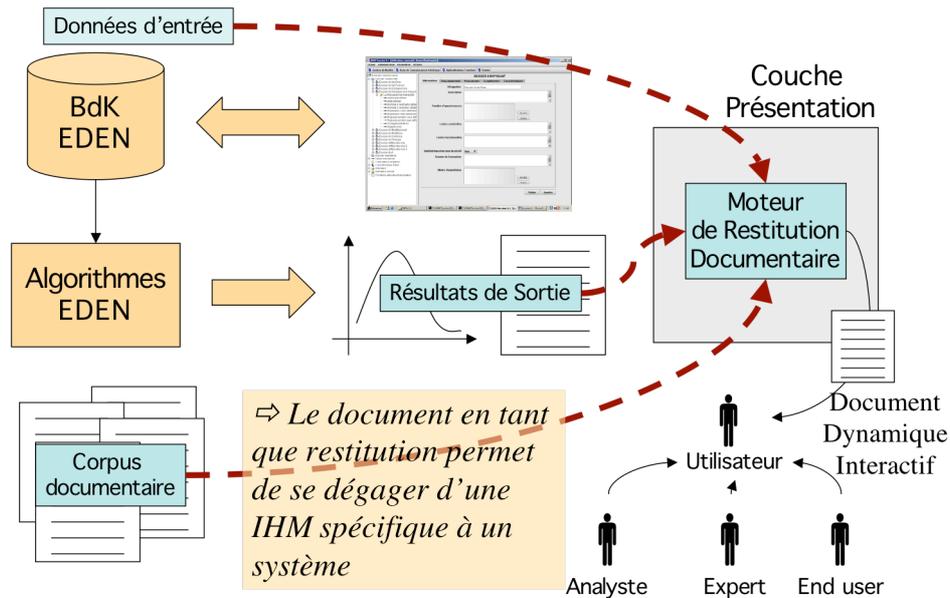


Figure 115 – Association du corpus documentaire et du système à base de connaissances (en tant que donnée d'entrée) pour produire des documents dynamiques interactifs constituant l'interface homme informations

Le moteur de restitution documentaire a pour vocation de calculer des documents dynamiques interactifs se posant en alternative des interfaces homme machine actuellement mises en œuvre. Nous allons illustrer 2 aspects relatifs à ce moteur :

- La production de document multimodaux,
- La modalité graphique.

### 3.2.a Document multimodal

La séparation entre contenu et mise en forme inhérente au document numérique structuré permet à partir d'une source de données unique de produire un éventail de visualisations de ce document. Au-delà de produire différentes formes de visualisation, il est possible d'une part de les associer au sein d'un même document mais également de les mettre en relation par le biais de mécanismes interactifs. Dans le cas qui nous occupe ici, nous avons ainsi produit à partir des données de la base de connaissances ou des résultats de sortie différents documents multimodaux intégrant à la fois une forme de représentation graphique et textuelle de la même information (cf. Figure 116).

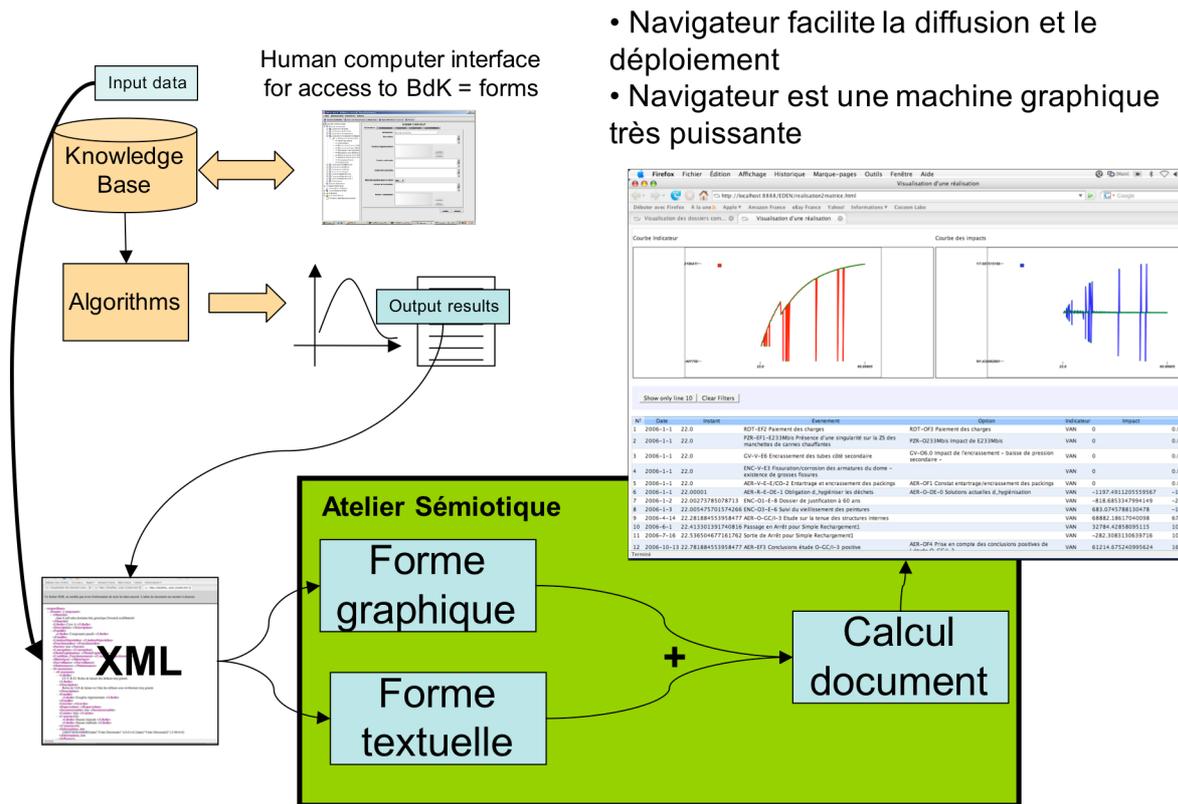


Figure 116 Processus de génération d'un document multimodal

La Figure 117 offre un aperçu d'un document produit permettant une visualisation d'un indicateur technico-économique. Cet indicateur est ainsi présenté dans 3 espaces de représentation :

- Un graphique à 2 dimensions qui permet de visualiser l'évolution de l'indicateur,
- Un graphique à 2 dimensions qui traduit les variations de l'indicateur,
- Un tableau de valeurs interactif.

La génération des graphiques à 2 dimensions s'appuie sur une bibliothèque spécifique intégrée à notre plateforme dédiée à la visualisation de graphiques. Le tableau s'appuie sur un widget disponible dans la boîte à outil dojo là encore intégré à notre plateforme sous la forme d'un dispositif.

Le document produit est alors un document multimodal dont les différentes modalités sont en relation via un mécanisme interactif. Ainsi, l'interaction avec une modalité provoque une répercussion sur les autres. Dans le cas présent, le fait de pointer et de cliquer sur un point d'une courbe va directement permettre d'accéder à sa représentation textuelle dans le tableau de valeur. On peut ici souligner tout l'intérêt de produire les images dans un format vectoriel puisqu'il est alors possible d'associer à toute partie de l'image un mécanisme interactif de

manière simple. Ainsi, la mise en relation se révèle intéressante dans l'analyse de la courbe car le tableau peut introduire des informations complémentaires relatives à la valeur représentée. La courbe peut donc constituer un moyen de navigation dans les données.

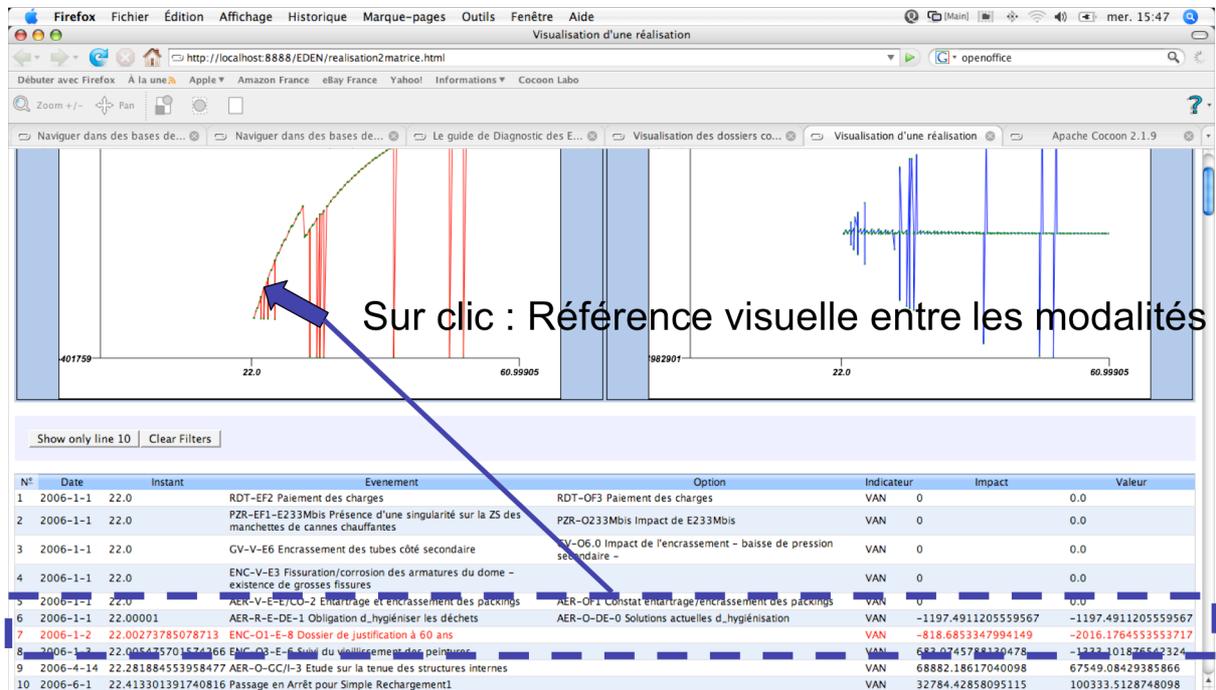


Figure 117 Visualisation d'un indicateur technico-économique sous la forme d'un document multimodal

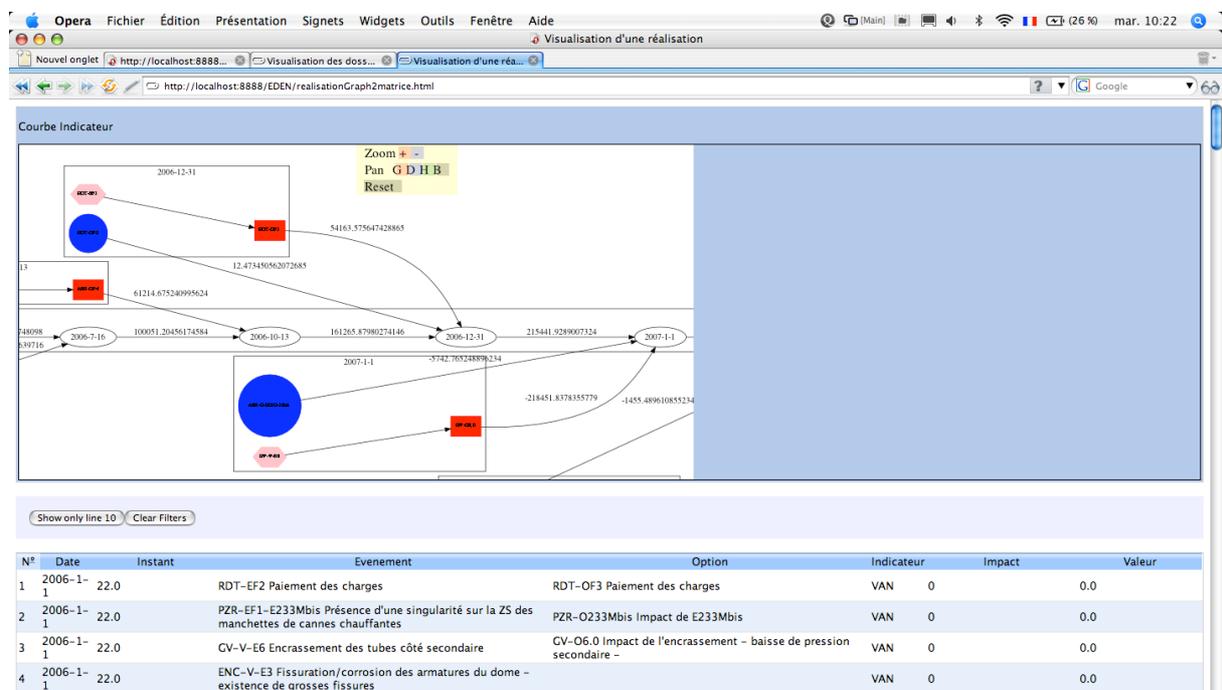


Figure 118 Document intégrant à la fois une visualisation de graphe et une visualisation tabulaire de la même information

Dans le même ordre d'idée, il est envisageable d'associer une visualisation de graphe à un tableau de données (cf. Figure 118), une arborescence à un document textuel (cf. Figure 119) ou encore graphique arborescence et tableau de données (cf. Figure 120). La mise en relation des différentes modalités est ainsi faite de manière interactive mais également visuellement. Les couleurs utilisées pour représenter les différentes briques d'information sont ainsi les mêmes dans les différents systèmes de représentation afin de faciliter la lecture et l'utilisation des documents en jouant sur la notion de références intermodales (Bestgen et Dupont 2003).

L'ensemble de ses documents même s'il nécessite une petite part de développement spécifique s'appuie sur un ensemble de bibliothèques génériques mettant à disposition des dispositifs de visualisation adaptés à nos architektes cibles. Les aspects de développement qui nécessitent encore une expertise informatique renvoient essentiellement à produire la combinaison de différents systèmes de visualisation.

The screenshot displays a web browser window titled "Naviguer dans des bases de connaissances". The address bar shows a URL: `http://localhost:8888/Sudoku/Navigation.html;jsessionid=4lwoanf444k0a?projectName=eden`. The browser tabs include "Démarrage", "Dernières nouvelles", "Passages à niveaux", and "Portail de navigation...".

The application interface has a top navigation bar with tabs: "Composants", "Réalisation Type", "Résultats", and "Deconnexion". Below this is a sub-tab "configuration" with a dropdown menu showing "mode : Dojo".

On the left side, there is a tree view titled "Les composants". The root node is expanded to show a list of components, each with a red square icon and a minus sign. The selected component is "Cpt - Alternateur tranche AF".

The right side of the interface shows the detailed view of the selected component. It includes the following information:

- Cpt - Alternateur tranche AF -**
- E - E1 : Détection de baisse d'isolement du bobinage stator par humidification**
- Typologie : V
- Tranche : A
- Contextes : Tous
- Probabilité d'occurrence :  $P1(t) =$
- Opt - O1.0 :**
- Mise en oeuvre : Sans objet
- Indispo : VD : 0 VP : j ASR : j
- Cout :
- Delta Cout Maintenance : 0
- Impact Puissance : Aucun
- Impacts sur les lois des evenements : Augmentation de P1(t). Intégrée dans P2(t)
- Opt - OC1.3 : Remplacement de barres en VD ou en VP Barres d'entrefer (EF) -**
- Mise en oeuvre : Arrêt programmé suivant l'événement
- Indispo : VD : 0 VP : j ASR : j
- Cout :
- Delta Cout Maintenance :
- Impact Puissance : Aucun
- Impacts sur les lois des evenements : Aucun
- Opt - OC1.3 : Remplacement de barres en VD ou en VP Barres fond d'encoche (FE) -**

The status bar at the bottom left of the application window shows "Terminé".

Figure 119 Une visualisation arborescente et synthétique est intégrée à droite du document textuel pour supporter la navigation

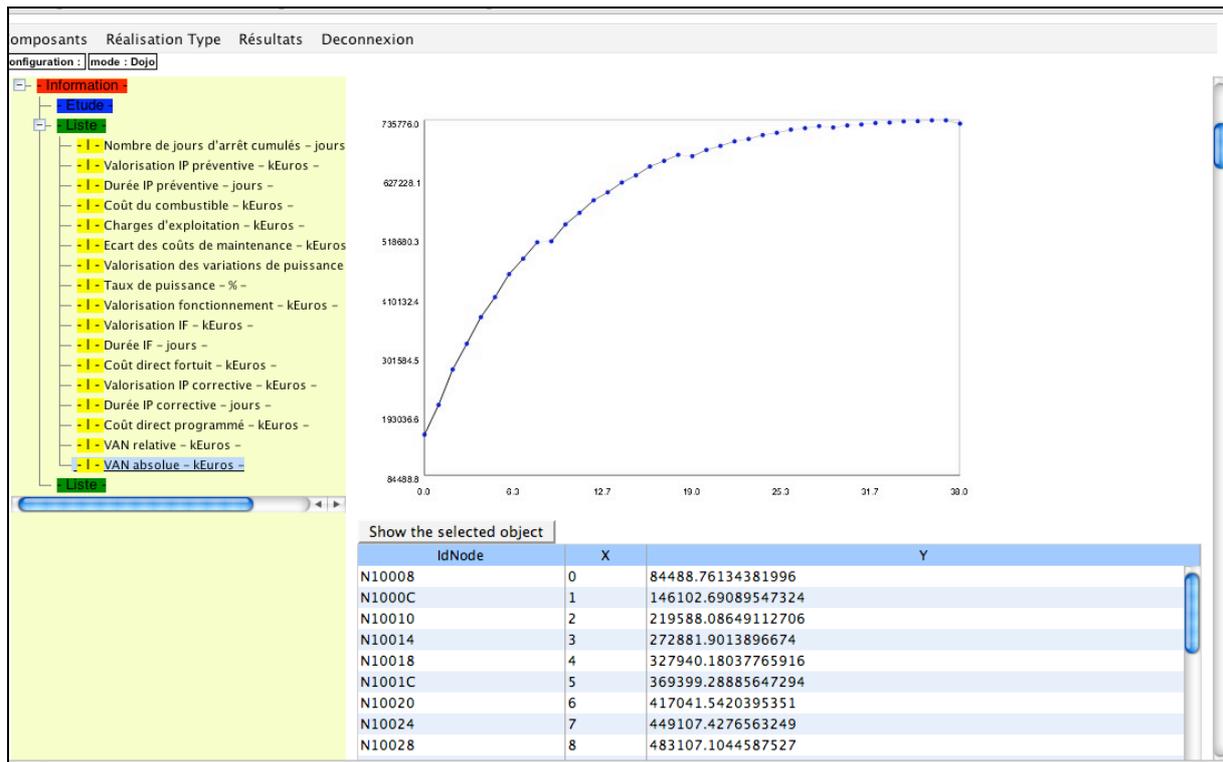


Figure 120 Document multimodal intégrant 3 formes de représentation : arborescence interactive, graphique et tableau de données

### 3.2.b La modalité graphique

Dans ce projet, la visualisation de graphe proposée a été introduite pour offrir un moyen de visualisation sur le contenu de la base de connaissance d'une part et sur les résultats de simulations d'autre part. Ces deux cas d'applications correspondent à une problématique de visualisation d'une information essentiellement relationnelle. La modalité graphique a été introduite pour faciliter la lecture des structures complexes de modélisation ou encore de données de simulation.

Nous avons naturellement adopté la visualisation de graphe (cf. Chapitre 5 -1.1. page 103) tout en respectant la philosophie de notre approche (cf. Chapitre 10 -2 - 180) inspirée de la sémiologie graphique (cf. Chapitre 5 -2 - page 116). L'ambition est à travers cette métaphore graphique que constitue le graphe, de proposer une image permettant de traiter l'information contenue dans la base de connaissance ou liée aux résultats de sortie. Un aperçu du type de visualisation produite est présentée sur la Figure 121 et la Figure 122.

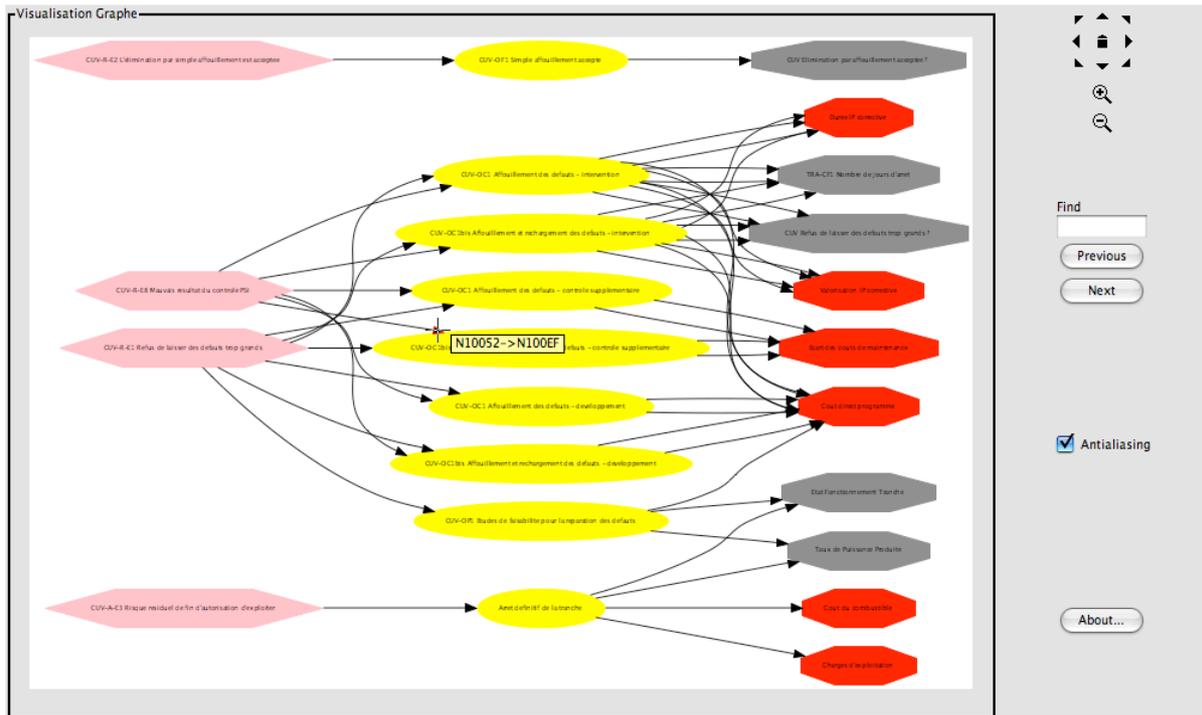


Figure 121 Visualisation sous forme de graphe du contenu de la base de connaissances

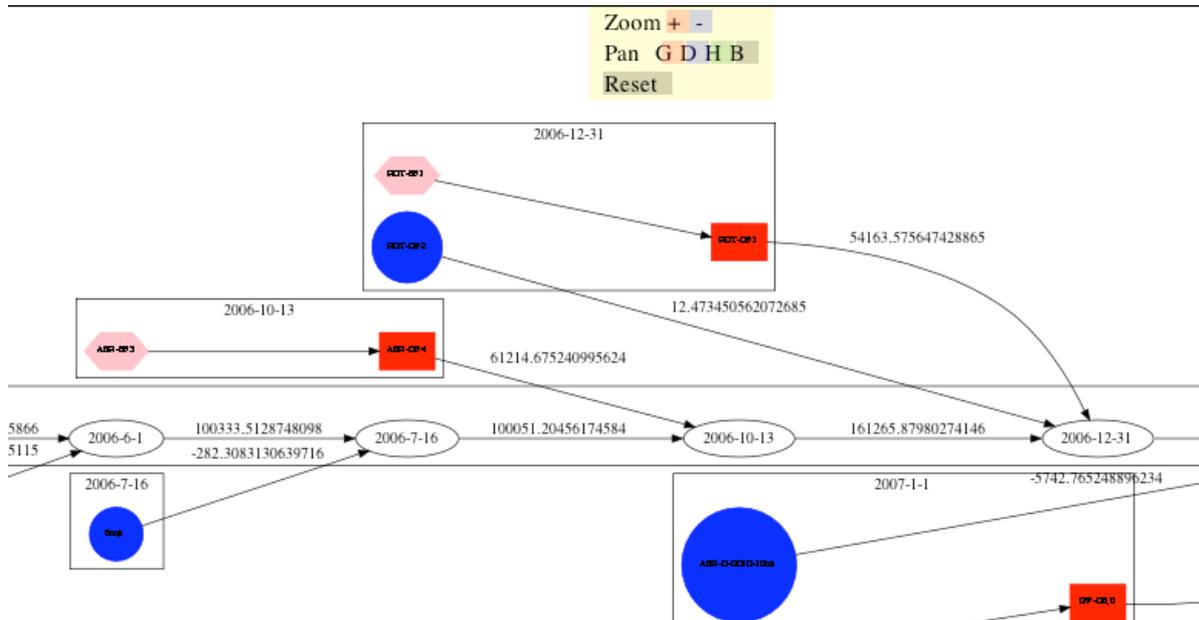


Figure 122 Visualisation sous forme de graphes de résultats de sortie

Techniquement, la production des visualisations graphiques s'appuie sur des images vectorielles codées en Scalable Vector Graphics (SVG). Sur la base de ce standard, l'image est enregistrée sous la forme d'une description au sein d'un document structuré. Chaque élément qui constitue l'image va alors correspondre à un élément balisé. Du fait du caractère

vectériel de l'image, il est possible de zoomer et de se déplacer sur l'image sans perte de qualité. Dans notre maquette nous avons en particulier utilisé l'aplet Java ZGRViewer (<http://zvtm.sourceforge.net/zgrviewer.html>) qui permet une navigation souple et rapide dans l'image. De plus, l'ensemble des éléments composant l'image peuvent devenir interactif et être mis en relation avec d'autres ressources informationnelles. Un scénario pourrait être, par le biais d'une visualisation graphique correspondant à la modélisation formelle, de pouvoir accéder directement aux documents de capitalisation expliquant cette modélisation. La forme documentaire peut représenter une représentation de l'information plus adaptée à un non-spécialiste du système.

La sémiologie graphique de Bertin met également l'accent sur la production d'une image monosémique. Une image monosémique se caractérise par le fait qu'elle ne doit permettre qu'une seule interprétation et donc que la sémantique graphique est clairement identifiée. Dans ce cas d'application, nous avons défini cette sémantique graphique de manière supervisée (cf. Chapitre 10 -2.6.b page 209). Ceci étant, elle aurait pu être définie dynamiquement par une analyse du contenu à représenter ou encore par l'utilisateur à travers une interface masquant le besoin de programmation. On lui donnerait alors la possibilité de modifier la représentation graphique par exemple pour mettre en valeur certaines informations sur lesquelles il veut se concentrer, tout en l'accompagnant dans le paramétrage afin de conserver une image monosémique.

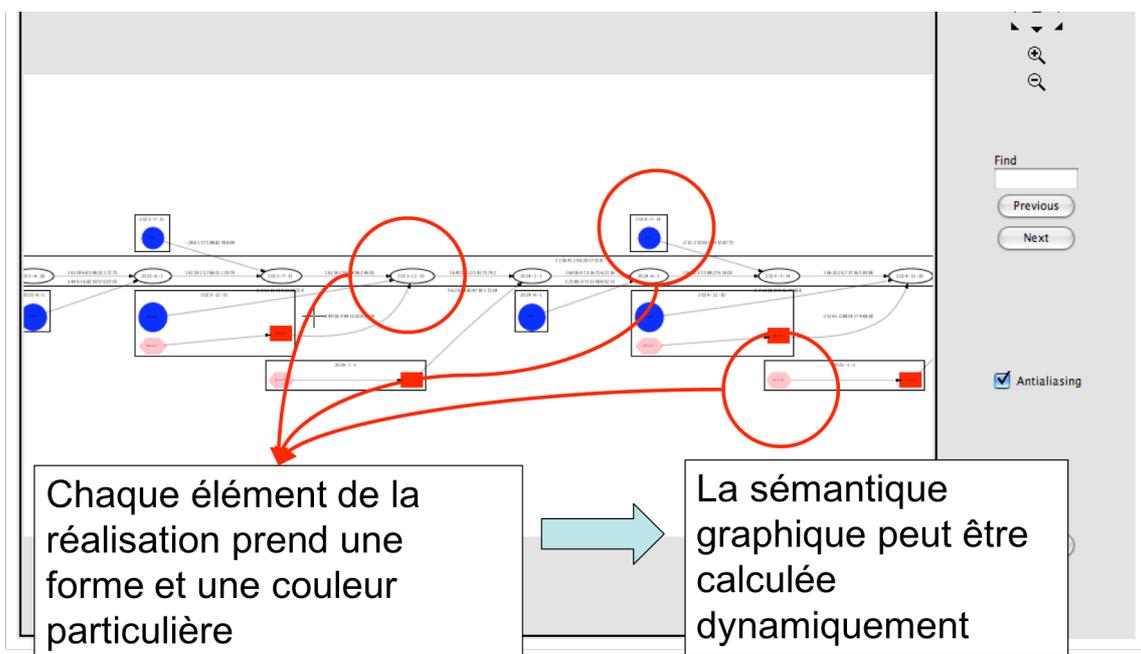


Figure 123 Sémantique visuelle dans la génération de graphe

### 3.3. Retour d'expérience

Dans le cadre de cette expérimentation, nous ne pouvons prévaloir d'un retour d'expérience aussi riche que dans le cas d'application des guides de diagnostic présenté précédemment. En effet, bien que nous ayons proposé, conçu et développé un ensemble de solutions et de maquettes, nous n'avons pu procéder à une évaluation aussi importante du fait de contraintes inhérentes au projet en lui même (disponibilité des personnels concernés, organisation du projet, etc...) et certainement par manque de temps. Les éléments décrits ici constituent donc une appréciation générale rédigée sur la base de discussions informelles ou résultant de présentations des maquettes. Nous allons également souligner certains éléments liés à la solution proposée qui offre des réponses à des problèmes bien identifiés relatifs à ce cas d'application.

Tout d'abord, notre approche constitue une solution légère et flexible de définition d'interface. L'éventail de proposition de restitution qui se distingue à la fois sur la forme (modalité textuelle, graphique ...) et sur le moyen d'accès (un éventail d'architextes) illustre la richesse de possibilité de restitution sans engager un effort de développement trop important. Le développement est alors moins coûteux notamment du fait que l'on s'éloigne d'une solution basée sur un client lourd tel que c'était le cas auparavant. La flexibilité de la solution évite de s'enfermer dans une seule solution de mise en œuvre. On identifie bien qu'avec un peu plus de travail, il aurait été possible de restituer les contenus dans des architextes très courants tel que Microsoft Word ou Microsoft Excel. Pour des raisons de facilité mais également une volonté au moins dans un premier temps de s'appuyer des technologies libres, nous avons produit par exemple des résultats accessibles via la suite bureautique OpenOffice ou encore l'éditeur de carte conceptuelle Freemind.

Dans le cadre de ce projet, la saisie des informations dans les bases représente une tâche assez fastidieuse du fait du caractère procédural de la saisie amené par la mise en œuvre de formulaires. Cette tâche est d'ailleurs rendue peu agréable, étant donnée la lourdeur de l'interface mise en place qui induit des temps de réponse pouvant se révéler longs. Là encore notre approche basée sur le document structuré semble représenter une solution plus légère sur le plan de la conception et du développement<sup>5</sup> ainsi que sur le plan de l'utilisation. Le caractère structuré du document permet de conserver une gestion de la cohérence des

---

<sup>5</sup> Créer un formulaire adapté à un utilisateur et une activité précise représente un coût de développement important.

données. Cette gestion de cohérence sur le document n'est par contre pas aussi bloquante (cf. Chapitre 8 -2.1. page 154) que dans la mise en œuvre de formulaire. Nous avons en effet mis l'accent sur le fait qu'il était possible de laisser durant la rédaction du document structuré certaines parties du document incohérentes ou encore de laisser temporairement certaines informations indispensables non renseignées (cf. Chapitre 9 -2.1.c page 164). On passe donc d'une activité de saisie dans un formulaire à une activité plus proche de la rédaction d'un document. La rédaction peut se faire par un éditeur spécialisé avec une visualisation en code source telle que nous avons pu l'évoquer (cf. 3.1.a page 244). La souplesse de la solution permet également d'offrir des solutions de saisie basées sur d'autres architectures comme nous avons pu l'évoquer au 3.1.b page 248. La visualisation sous la forme d'un code source montrant de manière concrète les éléments structuraux (les balises) du document est alors traduite dans une mise en forme matérielle particulière.

Le passage à la rédaction d'un document plutôt que la saisie dans une base de données laisse entrevoir une perspective intéressante dans ce cas d'application en particulier. En effet, ce derniers se caractérise par une méthodologie de capitalisation qui intègre, à la fois, la production de documents écrits et la saisie dans une base de données formelle (cf. Chapitre 1 - 3.2.b page 46). Nous avons mis l'accent sur le problème du caractère disjoint des informations à la fois inscrites dans les documents et la base de connaissances (cf. Chapitre 8 -2.2. page 156). Ceci amène en premier lieu à une double saisie de beaucoup d'informations puisque nous avons pu mettre l'accent sur le caractère très structuré de certains documents ou parties de documents et le fait qu'il est envisageable d'extraire et exporter ces contenus directement dans une base de données (cf. Chapitre 8 -1.1. page 148). La double saisie peut également introduire des incohérences entre documents et base de données dues éventuellement à des erreurs de frappe ou au processus de traduction de la note de capitalisation à la modélisation formelle qui induit nécessairement une phase d'interprétation. On voit ici que notre approche fondée sur la mise en œuvre du document structuré pourrait supporter l'ensemble de la démarche et du travail de l'analyste dans ce cas d'application.

Les notes de capitalisation pourraient ainsi être rédigées sous la forme de documents structurés sur la base d'un modèle de document permettant d'articuler des contenus peu structurés, pouvant être plus verbeux, avec des contenus intégrant une structuration forte. Cette structuration sera alors en relation avec le modèle de structuration formelle de la base de connaissances informatique mise en œuvre, permettant d'exporter les informations contenues dans la note directement en base. On procéderait alors à la mise en place d'un lien fort entre les documents et la base de données formelle qui pourrait au-delà de la capitalisation être

exploitée en restitution. Ceci pourrait permettre d'accéder directement aux documents à partir de base informatique, et inversement les deux formes de représentation pouvant s'enrichir l'une avec l'autre. Ce lien pourrait également être valorisé dans l'analyse des résultats de sortie puisqu'à ce niveau, pouvoir visualiser directement la partie de la note correspondant à une information pertinente dans les résultats pourrait très largement accompagner l'analyse. En particulier, on peut imaginer d'assister la construction des documents de synthèse qui présentent les résultats au travers de ce rapprochement entre documents et base de données. Ceci se matérialiserait par exemple par la possibilité d'importer dans un nouveau document des fragments des notes de capitalisation en relation avec les données que l'on souhaite expliquer plutôt que de procéder à de multiples « copier coller » qui constituent pour une part une solution artisanale sans compter que l'on ne conserve pas le lien entre donnée d'entrée et de sortie rendant la représentation de l'information à la fois moins précise et moins pérenne.

## Chapitre 12 - Discussion et Perspectives

Sur la base de la mise en œuvre présentée dans le chapitre précédent, nous pouvons dresser un bilan plus général de notre approche et de sa mise en application. Il s'agit de discuter un certain nombre de points en rapport direct avec notre proposition au-delà des éléments de retour d'expérience que nous avons déjà présentés au Chapitre 11 -2.3. page 239 et au Chapitre 11 -3.3. page 259.

### 1 - Discussion

Notre proposition représente une approche qui couvre le cycle de vie du document structuré de sa construction à sa restitution. Les points dont nous souhaitons discuter ici sont les suivants :

- La construction des documents,
- L'instrumentation de la démarche,
- La pérennité des informations inscrites,
- L'autonomie face à la réalisation d'un document.

#### 1.1. La construction des documents

Notre positionnement est de travailler en amont de la construction des documents et non sur une analyse a posteriori de documents préexistants. Plutôt que de subir un corpus documentaire de plus en plus grand, pouvant se révéler difficile à exploiter, de par son caractère inadapté à un traitement informatique, il s'agit d'offrir les moyens à nos utilisateurs, nos auteurs, de créer l'indexation – au sens repérage et identification de briques élémentaires dans le document manipulable par un traitement informatique - dès la création de ces documents. Cette indexation est fondée sur la définition de langage de structuration du contenu.

##### *1.1.a Le document structuré comme espace de définition d'un langage*

Notre expérience de formalisation au travers de documents structurés va en ce sens. L'objectif était bien de mettre l'analyste dans une position de création d'un document plus facilement

valorisable par un traitement informatique présentant une indexation fine de l'information. Ceci est passé par une sensibilisation de l'analyste aux techniques en relation avec la production des documents structurés afin de lui montrer les possibilités offertes à l'issue de la rédaction de tels documents. L'expérience aura permis de montrer l'intérêt d'exploiter le document structuré dans la démarche même de capitalisation de connaissances. Il constitue un instrument souple de structuration de l'information permettant de faire évoluer, de concert, structure et contenu sous un mode déclaratif et de faire émerger un modèle, un langage de structuration de l'information. Il n'est alors plus nécessaire d'anticiper sur le langage de structuration de l'information. Ceci est rendu possible par une démarche ne définissant pas de modèle de document (DTD , Xschema) a priori, celui-ci étant calculé a posteriori.

Le modèle de document recense alors l'ensemble des éléments structuraux introduits ainsi que leur organisation. Ceci va correspondre respectivement à la dimension lexicale et à la dimension syntaxique d'un langage qui permet de délimiter le contenu du document en cours de rédaction. Ce modèle, dès lors qu'il est stabilisé, peut alors conditionner la réalisation d'un ensemble de documents partageant un même objectif. Il peut également être la base du développement d'un système informatique plus formel.

#### *1.1.b Le document structuré comme espace de conception*

Le document structuré correspond donc à un réel espace de conception pour la définition d'un langage de structuration et cela au fur et à mesure de l'identification des informations résultant du processus de capitalisation. Le document structuré est support de l'inscription. Il conditionne et accompagne également l'écriture de cette inscription.

L'analyste a ainsi reconnu qu'avec sa montée en expertise sur les technologies de structuration et sa compréhension du document structuré, il modifierait aujourd'hui ses approches de capitalisation influençant ses techniques d'entretien.

#### *Le document structuré comme instrument de définition d'un langage*

Il s'agit donc de produire une information mieux structurée, mieux indexée et manipulable par la machine. L'objectif est de s'éloigner du mélange entre le fond et la forme auquel on aboutit par l'utilisation de logiciels Wysiwyg. Dans la réalisation de documents de capitalisation, le document structuré a permis de produire une ressource mixant des contenus codifiés à une granularité variable.

Nous avons expérimenté la possibilité pour un non informaticien de définir en autonomie son propre langage de structuration de manière empirique à mesure des informations qu'il

capitalise. Dans des problématiques où le contenu traduit une information relationnelle importante, on peut dès lors mieux maîtriser la cohérence des informations par la mise en œuvre de mécanisme de validation de documents. Le document structuré constitue d'ailleurs une solution permettant de faire évoluer plus facilement un document à la différence de la réalisation d'un document voué à l'impression. Notre méthode permet donc de produire une information manipulable par un traitement informatique mais sans pour autant être trop contraint par la définition de la logique de calcul (l'algorithme).

#### *Le document structuré comme support à la coopération*

A partir de là, bien que nous ayons laissé une grande liberté à l'analyste, le document structuré constitue tout de même un mode de représentation de l'information offrant de larges possibilités de manipulation par les informaticiens. En cela, il constitue un réel support de coopération dans la conception d'une ressource informationnelle à grande valeur ajoutée et s'associe tout à fait à la catégorie des documents pour l'action (cf. Chapitre 3 -3 - page 77). Ainsi, dans notre expérimentation alors même que le document structuré n'était pas stabilisé (en termes de structure et de contenu), il jouait déjà un rôle dans le réseau d'acteurs ; en effet, les experts l'utilisaient – sous une forme qui occultait le balisage – pour analyser et valider le travail de formalisation et de capitalisation réalisé par l'analyste.

Il est toutefois important de souligner que le contexte technique dans lequel nous évoluons est certainement un facteur qui joue en faveur de notre approche. Les informations sont par nature très structurées et en particulier de manière hiérarchique. Nous proposons d'ailleurs de discuter quelques limites en relation avec ce mode de structuration.

#### *1.1.c Limites du standard de document structuré*

La mise en œuvre de documents structurés, en s'appuyant sur le standard XML en particulier, n'est toutefois pas sans introduire certaines limites. Les difficultés sont en premier lieu liées au caractère hiérarchique de l'information.

#### *Un mode de représentation précoordonné*

La représentation hiérarchique est un mode de représentation classique et très répandu auquel on prête une facilité de lecture et un caractère intuitif dans son interprétation. C'est d'ailleurs cette représentation hiérarchique qui constitue une force du document structuré puisqu'elle rend son traitement et sa manipulation de manière informatique plus facile et plus déterministe et plus facilement maîtrisable. Cette propriété constitue par contre un biais car il

s'agit d'un système de représentation précoordonné de l'information. L'imbrication de l'information inhérente à la représentation va nécessairement influencer et contraindre l'auteur dans la construction du document.

#### *Une dichotomie résultant de la hiérarchie*

Une autre contrainte correspond à la dichotomie qui résulte de la structuration hiérarchique dans les contenus délimités par les éléments structuraux. On est en effet face à une succession et une imbrication d'éléments structuraux qui ne permettent pas le chevauchement (overlap) lorsque l'on délimite des fragments de documents.

De nombreuses solutions peuvent répondre à ce problème de manière informatique mais pourraient rendre le document moins facile à construire et en particulier à rédiger. Le problème dans notre approche est que nous nous adressons à un public de non informaticiens. Dans la rédaction des documents structurés, l'imbrication et le non recouvrement des éléments structuraux constitue un facteur de lisibilité et la saisie des documents.

#### *Sémantique inhérente à la hiérarchie*

Une seconde difficulté est liée à la sémantique de l'imbrication des éléments structuraux. Une des forces du document structuré est liée au fait que celui-ci est lisible dans une certaine mesure, à la fois par un humain et une machine. Il laisse en effet apparaître de manière immédiate la structuration contrairement à une base de donnée qui « disperse » l'information dans différentes tables. Ainsi, pour obtenir la vision apportée par le document structuré dans un système de gestion de bases de données, il est nécessaire d'exécuter un ensemble de requête. Cette force que présente le document numérique structuré laisse entrevoir le sens des contenus, par le biais des libellés des éléments structuraux. On identifie également la structure par l'imbrication mais pas nécessairement la sémantique associée à celle-ci. L'imbrication peut en effet traduire une sémantique importante qui n'est pas nécessairement explicite. Nous avons pu en donner des exemples au Chapitre 10 -2.3.b page 187 qui justifie la définition de notre modèle de domaine, ce dernier ayant pour vocation de rendre le maximum d'informations explicites. Ainsi, bien que nous exploitons le document structuré notamment parce qu'il facilite le développement d'une structuration en extension, il s'agit également lors de la création du document de bien inciter les auteurs à documenter cette imbrication en précisant l'intention associée à celle-ci lors de sa création (Nanard et Nanard 2005).

Notons toutefois que dans nos mises en œuvre, l'analyste non-informaticien a eu naturellement tendance à introduire des éléments structuraux supplémentaires spécifiant d'une certaine manière les liens d'imbrication. L'exemple présenté sur la Figure 124, illustre par exemple que sur le découpage du diagnostic, il a introduit dans la hiérarchie un élément « Etapes » au pluriel pour regrouper les éléments « Etape ». D'un point de vue structuration d'information, cet élément « Etapes » est superflu car le regroupement des éléments « Etape » peut être déduit par « toutes les éléments « Etape » qui appartiennent à une Phase » : il s'agit du type de traitement que pourrait opérer un informaticien. L'analyste pour sa part a eu tendance à transcrire ce type d'information car ceci lui paraissait plus clair dans sa formalisation en particulier pour lui-même.



Figure 124 Structuration faite par un analyste non informaticien face à une structuration informatique

### 1.1.d En restitution

#### *Flexibilité de la restitution*

Sur le plan de la restitution, notre approche ouvre un large champ de possibilités. Le document structuré se caractérise par une grande malléabilité. Il peut aisément être modifié à travers un traitement informatique. Par exemple, la forme du document produit en capitalisation ne constitue pas nécessairement la forme imposée de diffusion. Ceci représente une importante évolution par rapport à un document voué à l'impression. En particulier, cela

ouvre de nombreuses perspectives puisque l'on peut offrir un éventail de restitutions possibles du document « adaptées » à différentes tâches : le document est adaptatif. L'adaptation peut tout aussi bien impacter sur la forme ainsi que sur le fond. Elle permet de particulariser le document à différents supports permettant d'instrumenter l'activité de capitalisation de l'analyste mais également produire un document pour assister une tâche métier spécifique. Nous avons très largement discuté ce point dans la mise en œuvre (cf. Chapitre 11 - page 214).

#### *Création de documents monosémiques*

Notre démarche tend à produire des documents monosémiques. Ce caractère monosémique ne peut être considéré comme global : on ne peut prétendre que le contenu sera interprété d'une manière unique. Notre objectif était de passer outre la dilution de la sémantique du document en particulier dans la mise en forme. Ainsi, nos documents sont monosémiques jusqu'au niveau de granularité du document structuré produit. Il est alors possible d'explicitier la légende d'un document structuré restitué par notre approche. Cette légende, qui constitue le contrat de lecture et d'écriture dans le document, décrit la mise en relation des éléments structuraux contenus dans le document avec par exemple des règles de mise en forme. Ceci doit être un facteur pouvant assister la métacommunication – « méta » car elle est médiatisée par un médium que constitue le document - entre auteur du document et lecteur, l'ambition étant que ce dernier interprète correctement ce qu'il consulte. La construction de document monosémique constitue également un facteur de pérennité de l'information (cf. 1.3. page 270).

#### *Création de documents efficaces*

Dans notre proposition, le développement de la méthode de construction du document s'est également attachée à investir la philosophie de construction d'image proposée en sémiologie graphique (cf. Chapitre 5 -2 - page 116). Dans ce champ d'étude, une notion clé est l'*efficacité* d'une image dont résulte l'évaluation du caractère adapté ou non d'une variable visuelle pour représenter un type d'information dans une image. Nous avons évoqué que la sémiologie graphique est difficilement applicable à la construction d'un document. Aussi, notre méthode de construction essaie d'adapter cette technique de construction d'image au document. Au-delà de l'application au cas particulier du document, et tenant compte de notre contexte d'étude, cette méthode intègre également une prise en compte du contexte logiciel existant. Il ne s'agit pas de produire des documents uniquement lisibles et exploitables par

une unique interface que nous proposerions mais bien de pouvoir être en adéquation avec un architexte existant. Ceci passe par une analyse des architextes qui donne lieu à la définition d'un modèle d'architexte. C'est ce dernier qui intègre la notion d'efficacité. Il est construit à partir d'une évaluation des capacités de visualisation et d'interaction pour faire transiter un type d'information.

### *Le document comme cadre de conception d'IHM*

Cette prise en compte des infrastructures logicielles existantes nous amène à parler de la difficulté de concevoir une interface véhiculant des informations à un utilisateur (cf. Chapitre 4 -3 - page 99 pour la notion *d'interface homme information*). Notre positionnement est de promouvoir la forme documentaire comme une des interfaces les plus adaptée pour véhiculer des informations qu'elles soient destinées à l'impression ou à la consultation sur écran. Dans le cas présent, nous proposons une approche produisant des documents monosémiques et intégrant au moins pour partie la notion d'efficacité. Ces deux éléments ont pour vocation à la fois d'améliorer la métacommunication entre auteur et lecteur mais également le traitement de l'information à l'image de la sémiologie graphique. Nous avons évoqué le rapprochement entre document et interface homme machine au Chapitre 3 -4 - page 81. Notre étude appuie ce rapprochement. En outre, nous pensons que notre approche peut être une bonne méthode dans le développement des interfaces homme machine vouées à la communication d'information. Notons toutefois qu'il s'agirait de l'étendre et la compléter en particulier concernant la dimension interaction. Dans la description de notre approche, nous nous sommes essentiellement penchés sur l'utilisation des moyens de visualisation d'un architexte. Les moyens d'interaction n'ont été que peu abordés bien que nous ayons commencé à explorer leur prise en compte. Ceci constitue une voie d'exploration

### *Difficultés de restitution*

Tous les éléments exposés ici ne sont pas sans poser de difficultés. Les apports que nous venons d'exposer se font au prix du développement d'une instrumentation même si nous ne cherchons pas à proposer un outil global. Une autre difficulté est liée à la personnalisation de la restitution. Au stade de notre étude, il est clair que l'analyste n'est plus autonome dans la restitution des documents. Cette opération nécessite encore une expertise en informatique. Nous décrivons en perspective les grandes lignes de développement d'un Atelier sémiotique permettant de restituer à l'analyste une partie de cette autonomie tel qu'il la possédait dans la réalisation de documents via un logiciel Wysiwyg. Le résultat potentiel de cet atelier n'est pas

sans contre partie : il s'agira d'accepter certaines concessions. En particulier, le fait de lui proposer par exemple un ensemble de composants standardisés de visualisation et d'interaction complétés d'un accompagnement dans la construction de documents monosémiques et efficaces, tendra sur certains aspects vers une standardisation des documents produits. Des versions plus spécifiques ou plus adaptées pourront être développées mais en faisant appel à un informaticien.

## 1.2. L'instrumentation de la démarche

Comme nous avons pu le voir, notre proposition véhicule une instrumentation importante permettant de supporter la construction des documents. Notre étude a permis de souligner plusieurs points. Tout d'abord, il existe peu de solutions réellement efficaces, adaptées à un large public, pour construire et saisir des documents structurés. De plus, de nombreux logiciels informatiques bien que fondés sur le concept de WYSIWYG, sont exploités en WYSIWYM (cf. Chapitre 8 -1.3. page 152). Enfin, il semble difficile de vouloir bouleverser l'existant, sans compter que cela pourrait faire naître plus de difficultés que d'apports. Ceci nous a conduit à pouvoir être en accord avec les infrastructures logicielles existantes tout en n'écartant pas la possibilité de proposer de nouvelles solutions à condition de bien conduire leur déploiement.

De ce constat, notre orientation a très vite été de s'écarter d'une solution logicielle « tout en un » supportant la globalité de notre méthode. Il est évidemment confortable de pouvoir créer une solution spécifiquement adaptée à un problème. Nous n'avons pas opté pour cette orientation car cela nous semble complexe de la faire évoluer dans le temps. De plus, ceci ne constitue pas selon nous une orientation pérenne sachant que l'activité que l'on va outiller va nécessairement être modifiée du fait de l'introduction des outils (Béguin et Rabardel 2000). Nous y reviendrons au 1.3. page 270.

Notre positionnement est plutôt porté par le fait qu'il existe de nombreux outils adaptés à des utilisations et des tâches différentes. Ces outils représentent parfois les meilleures interfaces de lecture et d'écriture dans les systèmes d'inscription qu'ils proposent. Ils sont en effet spécifiquement en ce sens. Des exemples sont l'éditeur de cartes conceptuelles Freemind ou encore le tableur Excel pour une feuille de calcul. Ces outils représentent des interfaces souples et ergonomiques dans l'étape d'écriture et de lecture. De plus, ils sont employés car ils sont disponibles et nos utilisateurs savent les utiliser. Par contre, ils sont beaucoup moins adaptés à la gestion des connaissances comme dans nos cas d'application ou plus globalement

à la gestion d'informations guidée par des concepts tel que le repérage, l'indexation, la cohérence et la pérennisation d'informations. Ainsi, plutôt que de vouloir les remplacer, nous pensons qu'il est plus judicieux d'être en mesure de s'interfacer avec eux tout en les adaptant à nos usages et problématiques métiers.

C'est avec cette philosophie qu'a été mené le développement du moteur de restitution documentaire. Son fonctionnement est fondé sur des mécanismes à un niveau d'abstraction assez important permettant de s'adapter à un existant, éventuellement une solution propriétaire, tout en n'y étant pas cloisonné. Les exemples de mise en œuvre décrits ainsi que les développements effectués ont permis de démontrer qu'il était possible de s'interfacer avec de nombreuses briques logiciels existants sans entraîner un coût de développement important.

Nous évoquerons ici que la normalisation actuellement en cours dans le domaine des suites bureautiques au travers des formats OpenDocument (<http://www.odfalliance.org/>) et OpenXML (<http://www.ecma-international.org/publications/standards/Ecma-376.htm>) ne peuvent que nous encourager à poursuivre dans cette voie. Cette logique d'ouverture, s'écartant de formats propriétaires fermés, laisse entrevoir de nouveaux axes de capitalisation sur la forme au-delà du fond. Il ne s'agit plus uniquement en gestion des connaissances de capitaliser des informations mais également de tracer l'information traduisant la correspondance entre les dispositifs de mise en forme exploités et ce qu'ils représentent sémantiquement pour les auteurs lorsqu'ils les emploient. Ceci représente d'ailleurs peut être une voie d'évolution pour les logiciels existants. Si l'on prend l'exemple du traitement de texte Microsoft Word mais c'est également le cas pour celui de la suite OpenOffice, une fonctionnalité simple serait d'être en mesure de documenter les styles de mise en forme au-delà de l'attribution d'un simple nom. Ceci s'inscrirait dans une logique d'annotations (Zacklad 2007).

### 1.3. La pérennité des informations inscrites et de l'instrumentation

Notre approche véhicule une volonté et un objectif de pérennité des informations transcrites ainsi que de l'instrumentation proposée. Cette pérennité est tout d'abord liée au choix du document structuré lui-même. Quel que soit le standard technologique vers lequel on se dirige, il traduit l'information sous une forme très simple dont la structuration peut être identifiée par un humain sans nécessité d'un système de lecture complexe et spécifique. Ceci représente une importante différence avec une base de données dont les fichiers ne peuvent

êtres dissociés du moteur de base de données ayant permis de la créer pour être lus<sup>6</sup>. Le document structuré se matérialise le plus souvent par un simple « fichier texte » ne nécessitant qu'un simple éditeur pour être lu de manière élémentaire.

La pérennité est également liée au fait de s'être écarté d'une solution logicielle globale qui viendrait remplacer les outils existants. Le moteur de restitution documentaire opère essentiellement des traductions d'un système de représentation vers un autre, d'un architecte vers un autre. Ce processus de traduction se caractérise par le fait qu'il est totalement explicite. On conserve ainsi une inscription de la sémantique des informations. Cette explicitation correspond lors de la construction du document (cf. Chapitre 10 -2.6. page 203), à la mise en relation du modèle de perception avec le modèle d'architecte. Cette mise en correspondance de modèles constitue le langage de traduction dans le système de visualisation d'un architecte cible. Il se matérialise dans le moteur de restitution documentaire par un document structuré qui peut être rendu visible, selon notre approche, afin de produire une légende. On tend ainsi à limiter la dilution du document (cf. Chapitre 3 -2.2. page 76) évoquée par Bachimont, lorsque celui-ci découle d'une succession de transformations. Les documents générés en sortie de notre moteur sont monosémiques, jusqu'à un certain niveau de lecture, et l'on peut les accompagner d'une légende correspondant au contrat de lecture de la mise en forme par exemple.

#### 1.4. L'autonomie de l'analyste face à la restitution et la construction d'un document

La description de notre approche que cela soit dans son développement ou dans sa mise en application laisse deviner que la réalisation d'un document résulte d'une activité coopérative mettant en jeu analyste, informaticien et experts.

##### *1.4.a L'activité de l'analyste*

Dans le passage au document structuré, l'analyste reste le pivot dans la définition de la représentation de l'information (cf. Chapitre 6 -1.1. page 131). C'était bien notre objectif que de le laisser en autonomie dans le développement d'une ressource qui serait plus facilement exploitable par un traitement informatique. Ceci souligne l'importance de capitaliser des

---

<sup>6</sup> Les fonctions d'import / export permettent d'extraire le contenu d'une base de donnée. Toutefois, les informations seront toutefois toujours dissociés dans différentes tables sans compter qu'une telle extraction ne permet pas nécessairement d'obtenir le modèle conceptuel de données qui représente le contrat de lecture dans la base de données.

informations relatives à la structuration elle-même. Ainsi, il s'agira de bien documenter les modèles de documents (Marcoux 2006) auxquels on a abouti pour conserver une trace de la sémantique inhérente aux balises ou de l'intention de l'auteur au moment de leur création (Nanard et Nanard 2005). Une difficulté est par contre en relation avec la restitution des documents produits (cf. Chapitre 10 -1.4. page 180). L'analyste perd une part importante de son autonomie dans la construction et la visualisation document. Lorsque ce dernier exploitait un logiciel Wysiwyg, il produisait la forme qu'il souhaitait tout en structurant son contenu (cf. Chapitre 8 -1.3. page 152). Au regard de la restitution des documents, l'expertise de l'informaticien reste encore, au stade de développement de notre approche, incontournable.

#### *1.4.b L'activité de l'informaticien*

Malgré les efforts de développement sur le moteur de restitution documentaire, celui-ci nécessite encore de nombreux rudiments en informatique dans son paramétrage pour aboutir à la restitution du document dans un architexte. De plus, l'informaticien est le seul en mesure de créer les bibliothèques de dispositifs de visualisation et d'interaction (cf. Chapitre 10 -2.5.c page 199) permettant de produire une visualisation en accord avec un architexte particulier. L'informaticien est également celui qui va saisir le modèle de perception ainsi que la mise en relation de ce modèle avec le modèle d'architexte pour aboutir à la construction d'un document. Il est encore difficile de totalement évincer son expertise sur cette tâche bien que le moteur de restitution documentaire intègre la philosophie de l'*Atelier Sémiotique* évoqué au Chapitre 10 -3 - page 211.

De nombreux efforts pourraient être faits sur la plate-forme actuelle pour proposer des interfaces de saisie accessibles à un non-informaticien. Tout ceci appuie bien que dans la situation actuelle l'analyste n'est pas en autonomie face au processus de restitution. Le moteur de restitution documentaire actuellement développé intègre les prémisses de développement de l'*Atelier Sémiotique* que nous imaginons sans en offrir à la facilité d'accès.

## 2 - Perspectives

Dans cette section, nous ne reprendrons pas les perspectives d'évolutions associées à nos cas d'application que nous avons décrit en particulier au Chapitre 11 -2.3. page 239 et au Chapitre 11 -3.3. page 259. Avant d'aborder la perspective de développement de l'*Atelier Sémiotique*, nous décrivons une piste d'évolution essentielle de notre approche. Elle est liée au fait qu'il conviendrait de mieux prendre en compte les architextes existants au travers d'une analyse plus fine des variables disponibles. Ainsi, nous nous sommes largement concentrés sur les

moyens de visualisation pour tenter d'adopter la sémiologie graphique de Bertin à la construction de documents « figés » bien qu'intégrant des mécanismes de navigation. Il faudrait également mieux valoriser les moyens d'interaction. Ceci permettrait de révéler d'autant plus le potentiel du document numérique structuré.

## 2.1. Un approfondissement sur les moyens d'interaction

Nous avons déjà évoqué que la notion de document comme un cadre de conception d'interface nous semble une piste prometteuse. Dans l'étude que nous avons présentée ici, nous nous sommes essentiellement concentrés sur les moyens de visualisation. Une voie complémentaire, dans laquelle nous nous sommes déjà engagées, est d'approfondir l'analyse des moyens d'interaction d'un architecte, pour développer des dispositifs interactifs et évaluer l'efficacité des moyens qu'ils proposent face à tel ou tel type d'information. Nous dissocions l'interaction ici selon deux axes.

Le premier axe est l'interaction avec le contenu et en particulier sa modification. Prenons l'exemple d'une fonction de filtrage. On imagine assez bien que des composantes qualitatives constituent des informations permettant de filtrer ou de regrouper des informations assez facilement. De même, une composante ordonnée pourrait être un paramètre tout à fait adéquat pour arranger le contenu d'un document.

Le second axe est l'interaction avec un document restitué. Si l'on y accède via un terminal numérique, on peut envisager d'en modifier l'aspect dynamiquement (cf. Annexe 2 - page 305, pour un exemple de mécanisme d'*augmentation en information*). À partir de là, on peut projeter d'évaluer l'adéquation entre les mécanismes de masquage de fragments de documents (un exemple est le Titlepane de la bibliothèque Dojo présenté dans Annexe 1 - page 293) et certains types d'information.

## 2.2. Définition d'un atelier sémiotique et d'une méthode associée pour instrumenter la gestion des connaissances

Nous avons pu constater qu'il n'existe pas de réelle solution « tout en un » efficace pour la conception et la construction de documents structurés. Ceci oblige à se tourner vers une multiplicité d'outils. L'éventualité de développer une solution globale qui soit pérenne semble utopique. Ceci peut expliquer le recours à cette multiplicité d'outils ainsi qu'au développement de stratégies d'utilisation inédites par des utilisateurs parfois très inventifs.

Les développements que nous avons effectués dans notre étude ont permis d'aboutir à la réalisation d'un moteur de restitution documentaire mis en œuvre dans deux cas d'application réels. Cette confrontation a permis d'en évaluer l'intérêt mais également de mettre en évidence la nécessité d'une expertise en informatique dans son utilisation. Nous sommes tout à fait conscients qu'il conviendrait de faire évoluer cet outil sous la forme de ce que nous désignons comme un *Atelier sémiotique*. L'objectif à terme est de fournir aux analystes, dans leur activité de gestion des connaissances, un environnement leur permettant de composer une visualisation de leur document, avec une indexation de leur sémantique, sur le plan de la forme ainsi que du signe, sans connaissances avancées sur les technologies exploitées. Un aperçu de cet atelier a déjà été introduit au Chapitre 10 -3 - page 211.

### *2.2.a Description de l'atelier*

Cet atelier sémiotique aurait alors comme fonction, d'une part de permettre de produire une visualisation d'un document structuré selon un système sémiotique bien défini (déterministe et monosémique) en accord avec les possibilités de différents architectes. Les mécanismes de transformation mis en œuvre doivent être réversibles afin de pouvoir transformer un document pour qu'il puisse passer d'un architecte à un autre. Chaque architecte se transformerait alors en réel espace de construction et d'enrichissement du document, adapté à différentes activités et différents publics. Les architectes choisis constitueraient également une interface de diffusion et de lecture pour les utilisateurs, lecteurs, de ces documents. La Figure 125 illustre, dans les conditions où l'on part d'un document structuré produit par un analyste, le processus où s'insère notre atelier sémiotique. Il s'agit d'extraire à partir de ce document structuré le modèle de document (DTD, Xschema) correspondant qui va recenser l'ensemble des informations manipulables (les briques élémentaires) directement par la machine. De cette liste d'information, il s'agit de définir au sein de l'atelier sémiotique un modèle pour chaque architecte cible du document structuré (cf. Chapitre 10 -2.5. page 193). Le langage de représentation peut ensuite être appliqué sur une instance de document, pour produire une visualisation du document adaptée à une interface, un architecte cible. Ce langage correspondrait à la mise en relation du modèle de perception (cf. Chapitre 10 -2.4. page 189) et du modèle d'architecte (cf. Chapitre 10 -2.5. page 193) dans le cadre la construction du document (Chapitre 10 -2.6. page 203).

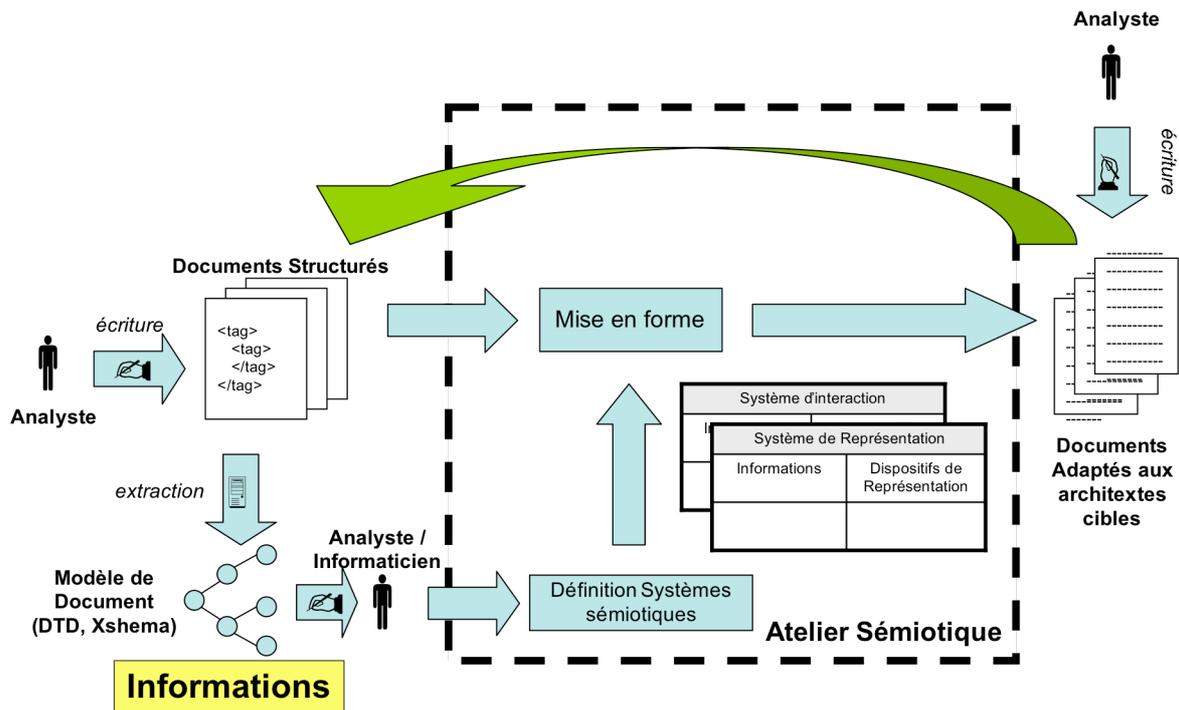


Figure 125 L'atelier sémiotique

### 2.2.b Intégration de l'atelier dans les différentes étapes d'un projet de gestion des connaissances

L'ensemble de cette étude s'est déroulé au sein d'une équipe pratiquant la gestion des connaissances en milieu industriel (cf. Chapitre 1 - page 21). Il semble évident que l'atelier sémiotique que nous venons de présenter aurait un intérêt certain à différentes phases des démarches de gestion de connaissances engagées. La Figure 126 illustre schématiquement le découpage d'un projet de gestion de connaissances dans notre version d'étude.

Dans l'étape d'*analyse*, notre proposition et en particulier l'atelier permettraient de donner un aperçu de ce que l'on est capable de réaliser au niveau de la capitalisation et la restitution de connaissances. En effet, l'atelier sémiotique représenterait un outil de prototypage rapide permettant plutôt que de discuter de concepts et de représentations hypothétiques, de proposer des exemples concrets d'orientation. Ceci serait rendu possible à la condition que cet atelier possède un ensemble de bibliothèques assez riche, permettant de s'interfacer avec de nombreux architectes et ainsi offrir un éventail large de modes de restitution possibles. Il pourrait ainsi permettre de proposer des démonstrations rapides, sur un jeu de données minimal, dans le cadre des réunions avec des acteurs comme les commanditaires, les experts

ou encore les utilisateurs, permettant d'alimenter la réflexion de chacun et inciter aux propositions.

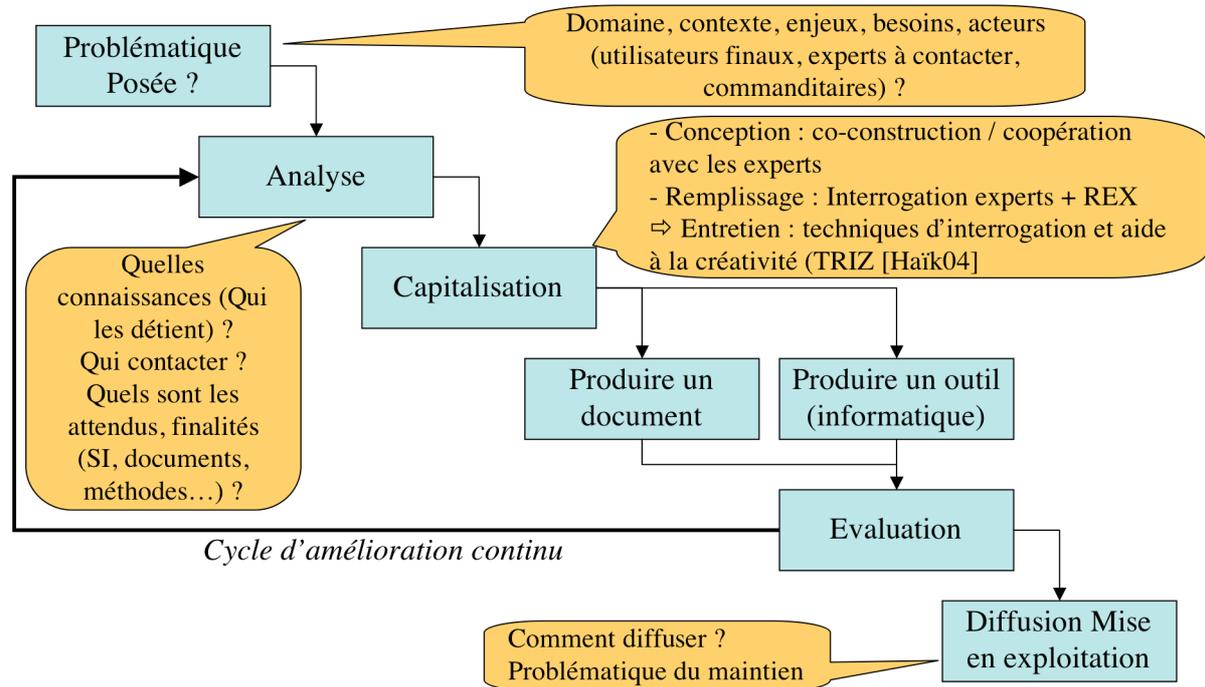


Figure 126 Les étapes d'un projet de gestion des connaissances

L'étape de *capitalisation* a abondamment alimenté notre étude. Nous avons donc déjà décrit tout l'intérêt de notre proposition. Elle offre une méthode et une instrumentation souples qui permettent de s'écarter du besoin de définir un modèle de structuration de l'information a priori. Celui-ci va émerger de la capitalisation elle-même. Notre approche peut également venir en complémentarité d'une solution de capitalisation existante. Son apport sera alors par exemple d'offrir des interfaces de saisie plus adaptées aux utilisateurs.

La *production d'un document ou d'un outil informatique* renvoie au fait que les projets de gestion de connaissances tendent généralement vers la conception et le développement d'un artefact. Celui-ci pourra outiller la capitalisation mais en premier lieu il s'attache à restituer les connaissances capitalisées sous une forme particulière. Tout comme la *capitalisation*, la *restitution* constitue une des problématiques sur lesquelles nous nous sommes concentrés (cf. Chapitre 2 -2 - page 52). Nous avons ainsi pu démontrer tout au long de l'étude et de nos mises en œuvre l'intérêt de notre approche et de son instrumentation. Nous fournissons une solution qui permet une restitution souple et flexible des informations capitalisées. L'approche de restitution s'attache à conserver au mieux la sémantique en produisant par exemple des documents monosémiques. L'atelier sémiotique pour sa part pourrait prendre le rôle d'un outil fédérateur dans le développement de solutions, constituant un environnement

de développement. Ceci pourrait faciliter la mise en relation de certains projets de gestion des connaissances connexes par exemple.

Au regard de l'*évaluation*, notre proposition n'engendre pas un apport direct. Nous pouvons toutefois constater que le fait de proposer une solution aussi souple de développement, permettant de construire avec les différents acteurs ne peut que permettre de converger plus rapidement vers une solution adaptée.

Enfin, les apports évoqués dans le cadre de la restitution joue également un rôle dans la *diffusion et le maintien de la ressource produite*. Nous évoquerons également les facteurs de pérennité de notre approche, sur le plan méthodologique et instrumentale, représente une propriété essentielle dans le cadre d'un déploiement.

### *2.2.c Valorisation de l'approche dans d'autres activités*

En prolongement de la gestion des connaissances, il est possible de dresser une perspective plus large de l'application de notre approche ainsi que du potentiel de l'Atelier Sémiotique. A travers ce mémoire, on identifie assez bien que notre proposition n'est *pas uniquement centré sur l'activité de gestion des connaissances*. Il s'agit d'une solution centrée sur *la construction et la restitution de documents structurés*. Plus largement, cette approche peut trouver sa place dans la gestion d'informations à un niveau plus général.

En particulier, le groupe de recherche dans lequel nous avons évolué, dans le cadre de cette étude, ne se compose pas uniquement d'une compétence en gestion des connaissances. Il s'agit d'un groupe multi-disciplinaire incluant également une compétence en traitement du signal ainsi qu'en automatique. Dans le cadre d'une action ponctuelle, nous avons été sollicités par la compétence automatique. Le problème posé était à partir de données scientifiques – dans le cas présent elles résultaient d'analyses sur des données de capteurs par le biais du logiciel de calcul scientifique *Scilab* (cf. <http://www.scilab.org/>) – de produire des documents pouvant être complétés par des personnels de terrain. Dans ce cadre, nous avons développé une maquette construite sur la base de notre proposition, qui a suscité un intérêt certain. Ce cas d'application, même s'il s'agirait de préciser les différences et similarités avec notre terrain d'étude, laisse entrevoir le potentiel du document structuré en général, de notre approche de construction et de restitution en particulier. Il souligne également l'intérêt du développement de l'Atelier Sémiotique dans des activités de Recherche et Développement où l'information, son inscription, sa manipulation et son traitement, sa diffusion et sa pérennisation jouent un rôle majeur.

## Conclusion

Une des motivations principales de ce travail de thèse est de contribuer à une amélioration de l'inscription d'information sur des supports numériques. Cette amélioration doit être conduite en particulier pour la rendre plus adaptée à un traitement informatique (que cela soit pour une exploitation par un algorithme ou une restitution destinée à un utilisateur). Nous nous sommes intéressés en particulier au cas du document qui joue un rôle prépondérant à la fois dans la construction des informations, ainsi que dans leur diffusion. C'est en particulier le cas dans l'activité de gestion des connaissances qui a été notre terrain d'étude, mais également dans bien d'autres activités d'entreprises et d'organisations.

L'observation permet de constater que la construction des documents est encore essentiellement produite par le biais de logiciels Wysiwyg (« What you see is what you get ») souvent basés sur l'utilisation formats propriétaires. Ces logiciels sont très largement employés car ils sont rendus disponibles par les directions informatiques, ce qui a permis aux utilisateurs de développer, au fil du temps, des habitudes d'utilisation en adaptant ces outils à leur activité. Nous avons d'ailleurs souligné que les utilisateurs exploitaient ces logiciels en Wysiwym (« What you see is what you mean ») au sens où, dans la construction du document, ils développent des langages de structuration et de représentation d'informations. Notons bien que ce développement de langages et la forme structurée que prend l'inscription est très lié au contexte technique dans lequel nous évoluons : l'information technique est par nature très structurée.

Cette pratique traduit le fait que le document constitue, dans sa rédaction, un espace de conception où un auteur développe sa créativité. On voit ici que le langage de structuration est conçu dans l'action et il ne semble pas possible de réellement pouvoir anticiper de façon détaillée sur celui-ci : on pourra, au mieux, faire une analyse a posteriori des documents pour en analyser la structure, le contenu et la mise en forme avec l'objectif d'identifier les langages définis par un auteur. Le problème posé par ce type de démarche est que le résultat produit sera nécessairement moins précis que si l'auteur avait produit l'indexation lui même en rendant explicite, par exemple, la sémantique liée à la mise en forme matérielle.

Ainsi, notre positionnement est de travailler en amont de la construction des documents plutôt que sur une analyse a posteriori. De là, notre choix s'est porté vers l'utilisation de document numérique structuré, le considérant comme un support d'inscription adapté au problème posé. Ce choix est justifié par plusieurs raisons. Une première est qu'il correspond à une représentation numérique de l'information aisément manipulable par un traitement informatique, tout en restant indépendant de l'outil ayant permis de le créer. Il est en effet très différent d'une base de donnée, qui offre des possibilités de traitements informatiques au moins identiques, mais qui n'est pas indépendante du moteur ayant permis de la créer. Une deuxième raison est que la réalisation d'un document structuré sous sa forme brute (un code source balisé), ne nécessite pas une expertise en informatique importante sous deux conditions. D'une part, il faut tout de même en expliquer les rudiments à un non-informaticien : ceci constitue une contrainte raisonnable. D'autre part, cette facilité de réalisation est liée au fait de ne pas lui imposer directement une logique forte de saisie tel qu'un modèle de document. Dans le cas contraire, on serait assez proche d'une interface de saisie de type formulaire avec toutes les difficultés qu'elles véhiculent (cf. Chapitre 8 -2.1. page 154). Une troisième raison est l'éventail de possibilités dans le cadre de la restitution du document aux utilisateurs. A partir d'un ou plusieurs documents structurés, il est en effet possible de produire un ensemble de visualisations, adaptées à une impression sur papier ou à une consultation sur écran, et dans ce dernier cas intégrant éventuellement de l'interactivité. Bien que le document structuré représente un choix séduisant au premier abord, il subsiste à notre sens 3 difficultés principales.

En premier lieu, il n'existe pas de réelle méthode de construction de documents structurés en dehors de celles proposées en conception logicielle. Cette dernière tendra à produire un modèle de document a priori, ce qui ne semble pas adapté à notre contexte d'étude. La seconde difficulté est liée à une des qualités du document structuré. Sa facilité de traitement permet, en effet, d'opérer de nombreuses modifications sur un document original. C'est alors la tentation de le considérer comme un « lego » qui risque de conduire à une dilution du contenu au fil des différentes opérations de traitement matérialisées par des algorithmes (cf. Chapitre 3 -2.2. page 76). Enfin, on constate la faible offre logicielle permettant d'accompagner la construction et la restitution de documents structurés adaptée à des non-informaticiens.

De ce constat, dans le cadre de cette thèse, la contribution scientifique de notre étude tient à deux aspects. Tout d'abord, nous avons tenté d'articuler 3 champs disciplinaires que sont :

l'ingénierie documentaire, l'ingénierie des interfaces homme machine et ce que nous nommons l'ingénierie de la forme, regroupant la visualisation d'information et la sémiologie graphique. Cette articulation nous amène d'une part à penser que la notion de document représente un bon cadre de développement pour guider la conception des interfaces homme machine (Parfouru et al. 2006). Nous avons également souligné que la sémiologie graphique pouvait guider la construction de documents d'un type particulier au prix d'une extension. Ce dernier point constitue à nos yeux notre deuxième contribution. Nous avons ainsi proposé une approche méthodologique et instrumentale de construction et de restitution de documents numériques structurés (Parfouru et al. 2007).

La méthode que nous proposons est assez largement inspirée de la méthode de construction d'image proposée par Bertin en Sémiologie graphique. Elle a été étendue pour répondre aux spécificités du document, mais également pour pouvoir tenir compte des contraintes réelles de terrain. Ce dernier point est lié au fait que nous ne souhaitons pas proposer une solution logicielle « tout-en-un » venant en substitution des infrastructures logicielles existantes. La réalisation d'une telle plate-forme nous semble utopique, sans compter que sa pérennité dans le temps serait très réduite. En effet, quel que soit l'outil introduit dans une activité, il aura tendance à modifier la tâche. L'utilisateur se l'approprie et construit des stratégies d'utilisation (schèmes) qui lui sont propres, le transformant en instrument (Béguin et Rabardel 2000) : ceci tend à privilégier des solutions légères et adaptables, voire même jetables, si l'on considère que les besoins peuvent évoluer rapidement. Nous avons privilégié d'être en mesure de produire une restitution documentaire compatible avec les architextes déployés afin de profiter des habitudes et stratégies d'utilisation déjà existantes. Il s'agit de ne pas imposer à l'utilisateur une nouvelle interface sur laquelle il lui faudrait reconstruire un ensemble de pratiques. Chaque architexte va alors poser des contraintes sur l'utilisation des variables visuelles qui vont devoir être prises en compte obligeant à adapter là encore la sémiologie graphique. Soulignons toutefois qu'il ne s'agit que d'un canevas qui trace les grandes lignes guidant la construction d'un document. Il reste de nombreux points à investiguer tels que la prise en compte des *moyens d'interaction* avec la notion d'efficacité (Chapitre 12 -2.1. page 273) ou encore la définition de *modèle éditorial de documents* (cf. Chapitre 10 -2.6. page 203) lorsque l'on souhaite associer plusieurs espaces de représentation.

Notre approche a été mise en œuvre dans le cadre de deux projets réels de gestion de connaissances au sein de EDF Recherche et Développement. Le premier correspond à une action de gestion de connaissances dans le domaine de l'hydraulique (cf. Chapitre 1 -3.2.a

page 43). Dans ce projet, nous avons appliqué l'ensemble de notre méthode et instrumenté la capitalisation et la restitution de connaissances (Grassaud et Parfouru 2007). Nous avons ainsi opéré un passage d'un document, produit via un traitement texte classique, qui devenait difficile à mettre à jour et à maintenir, vers une ressource numérique « vivante » et dynamique offrant de larges possibilités de restitution (Parfouru et al. 2006; Parfouru et al. 2007). Le second projet s'ancre dans le domaine du nucléaire (cf. Chapitre 1 -3.2.b page 46). L'objet était de compléter un système à base de connaissances formel, pour en faciliter l'accès à des utilisateurs n'ayant pas une expertise assez importante sur le système. Dans un premier temps, nous avons proposé une évolution des modes de restitution de l'information sous la forme de documents interactifs (Mahé et Parfouru 2006; Haïk et al. 2007). Dans un second temps, nous avons proposé une instrumentation de la capitalisation des informations alimentant la base de connaissance en s'appuyant sur des formes documentaires de restitution (Parfouru et Haïk 2007; Haïk et Parfouru 2008).

La philosophie de l'approche a également été investie dans deux autres projets d'EDF. Le premier renvoie à une évolution, menée par Benoît Ricard, des fonctions de restitution de résultats de diagnostics issus d'un outil existant (Ricard 2007). Dans ce projet, on a, en particulier, tenté de valoriser la forme documentaire de restitution des informations construites en considérant « *la contrainte d'exhaustivité de l'information présentée et l'ambition de tenir sur une page imprimée afin de préserver la simplicité de manipulation du résultat* ». Le deuxième projet se situe dans la continuité des guides de diagnostic. Il s'agit à la fois de reprendre les maquettes existantes pour une pré industrialisation de la solution, tout en améliorant certains mécanismes déjà développés. C'est dans cette démarche que s'insère le stage de Jérôme Gagnière dans le cadre de sa formation d'ingénieur (Gagnière 2008).

L'ensemble de ces mises en œuvre, nous ont permis de répondre à des problèmes bien identifiés dans notre contexte d'étude et d'obtenir une évaluation qualitative de notre approche. L'approche que nous proposons offre ainsi de nombreux apports que cela soit en capitalisation ou en restitution de connaissances dans le cadre de notre terrain d'étude (cf. Chapitre 11 -2.3. page 239 et Chapitre 11 -3.3. page 259). Dans le cadre de la capitalisation, cette recherche a bien montré la possibilité à un non-informaticien de construire des documents structurés directement en code source balisé au prix d'une explication de quelques rudiments en informatique. De plus, le document structuré s'est révélé un espace de conception (cf. Chapitre 12 -1.1.b page 263) adapté à cette activité de capitalisation à condition de ne pas imposer a priori sa structuration. Dans le cadre de la restitution, le

document structuré, sous entendu par notre approche, offre une grande flexibilité dans la restitution des informations (cf. Chapitre 12 -1.1.d page 266). Il est alors possible de décliner l'information selon un éventail de formes possibles construites en conservant la sémantique de la mise en forme, par exemple, et en gardant à l'esprit la notion d'efficacité définie par Bertin.

Plus largement, nous avons mis l'accent sur l'intérêt et les limites de la démarche au regard de la *construction de documents numériques structurés* (Chapitre 12 -1.1. page 262) ou encore la *pérennisation des informations inscrites* (Chapitre 12 -1.3. page 270). Il subsiste une difficulté importante au stade de notre étude, correspondant au caractère nécessaire d'une expertise en informatique dans l'application de notre démarche. Dans un projet de gestion des connaissances, par exemple, un analyste n'est plus en autonomie face au processus de restitution des informations comme il pouvait l'être en utilisant un logiciel Wysiwyg. En réponse à ce problème, nous avons tracé les grandes lignes de développement d'un *Atelier Sémiotique* utilisable par un non-informaticien, et permettant d'accompagner la construction et la restitution de documents structurés (cf. Chapitre 10 -3 - page 211 et Chapitre 12 -2.2. page 273).

## Liste des figures

Figure 1 Différentes vues relatives à l'activité de représentation de connaissances .....	22
Figure 2 Exemple de fichier CSV dont les informations sont délimitées par « , » et en regard son équivalent matriciel (source : Wikipedia).....	30
Figure 3 Exemple de fichier XML (source : Wikipedia).....	31
Figure 4 Un exemple de flux RSS (source : Wikipedia).....	32
Figure 5 Exemple de taxonomie (source : <a href="http://www.theses.ulaval.ca/2004/22242/ch02.html">www.theses.ulaval.ca/2004/22242/ch02.html</a> )....	37
Figure 6 Représentation graphique d'une ontologie.....	37
Figure 7 – Le guide de diagnostic des conduites forcées construit au travers du traitement de texte Microsoft Word .....	45
Figure 8 – Le guide d'aide au diagnostic des écluses construit via microsoft Visio et déployé via navigateur WEB .....	45
Figure 9 Vue schématique du contexte organisationnel et technique du projet.....	47
Figure 10 - Réseau d'acteurs dans lequel s'insère une activité de gestion des connaissances	50
Figure 11 Un exemple de formulaire.....	59
Figure 12 Exemple de code html et de sa visualisation .....	66
Figure 13 – Document structuré en XML sous différentes formes .....	67
Figure 14 Exemple de document structuré en Tex. (source : <a href="http://www.nanotechnologies.qc.ca/propos/linux/documentation/Redaction_documents_Latex/ch-structure_latex.html">http://www.nanotechnologies.qc.ca/propos/linux/documentation/Redaction_documents_Latex/ch-structure_latex.html</a> ).....	68
Figure 15 Document Tex mis en forme .....	68
Figure 16 Exemples de balisages généralistes .....	69
Figure 17 Quelques exemples de règles devant être respectées pour qu'un document XML soit bien formé .....	70
Figure 18 Illustration d'un modèle de document sous la forme d'une DTD (source : <a href="http://pagesperso-orange.fr/emmanuel.remy/XML/DTD/DTD.htm">http://pagesperso-orange.fr/emmanuel.remy/XML/DTD/DTD.htm</a> ) .....	71
Figure 19 Illustration de la validation d'un document XML (source : <a href="http://pagesperso-orange.fr/emmanuel.remy/XML/DTD/DTD.htm">http://pagesperso-orange.fr/emmanuel.remy/XML/DTD/DTD.htm</a> ) .....	71
Figure 20 Illustration de l'empilement des couches d'abstraction dans un système d'exploitation.....	84
Figure 21 Architecture Modèle Vue Controleur ( <a href="http://www.labo-sun.com/ressource-FR-articles-596-0-java-j2ee-introduction-a-j2ee.htm">http://www.labo-sun.com/ressource-FR-articles-596-0-java-j2ee-introduction-a-j2ee.htm</a> ).....	86
Figure 22 Modèle de communication de Shannon et Weaver (1948) .....	90

Figure 23	Modèle de communication de Jakobson (1961) .....	90
Figure 24	Le modèle de Riley (source : <a href="http://extension.missouri.edu/explore/comm/cm0109.htm">http://extension.missouri.edu/explore/comm/cm0109.htm</a> ) .....	92
Figure 25	La standardisation des icônes entre différentes applications Adobe (source : (Beaudoux 2004) ) .....	94
Figure 26	Courbe et tableau comme exemple de Data visualization.....	104
Figure 27	Exemple de visualisation de graphe et d'une carte conceptuelle .....	105
Figure 28	MétoMéta : un outil pédagogique de présentation de normes sous une forme graphique : une carte de méto ( <a href="http://www.mapageweb.umontreal.ca/turner/meta/francais/index.html">http://www.mapageweb.umontreal.ca/turner/meta/francais/index.html</a> ).....	105
Figure 29	(source : <a href="http://www.lucas.ch/">http://www.lucas.ch/</a> ).....	106
Figure 30	– La table périodique de méthodes de visualisation .....	107
Figure 31	TileBars (Hearst 1995), visualisation de la distribution des paragraphes pertinents dans un document (source : (Beaudoin-Lafon et Hascoët 2001)) .....	112
Figure 32	Liste indentée.....	113
Figure 33	Génération d'un graphe orienté basé la décomposition en niveaux d'un graphe..	113
Figure 34	Approche surfacique de visualisation (ex : tree-maps (Shneiderman 1992)) d'arbre ou l'imbrication des cadres traduit la hiérarchie et la profondeur à l'image d'une poupée russe (source : (Beaudoin-Lafon et Hascoët 2001)) .....	114
(Figure 35	Visualisation hyperbolique d'un arbre (source : <a href="http://ontology.univ-savoie.fr/condillac/en/activities/research/cs/">http://ontology.univ-savoie.fr/condillac/en/activities/research/cs/</a> ).....	114
Figure 36	Génération d'un graphe orienté avec un algorithme de « force direct placement ». .....	115
Figure 37	– Illustration des variables rétinienne dans le système graphique .....	120
Figure 38	Contexte de développement des guides de diagnostic .....	130
Figure 39	Contexte de capitalisation dans le cadre de la réalisation du guide de diagnostic	133
Figure 40	Les outils, interfaces et architextes exploités comme support à la coopération et la construction de la représentation des connaissances.....	135
Figure 41	Le point de vue métier.....	139
Figure 42	L'AMDEC.....	140
Figure 43	Le guide de diagnostic comme une articulation de points de vue .....	141
Figure 44	Contexte Socio technique dans le projet relatif à l'aide à la gestion d'actifs de production.....	143
Figure 45	Restitution documentaire dans le cadre d'un système à base de connaissances ...	146

Figure 46 Exemple de tableau mis en regard avec la modélisation résultante.....	149
Figure 47 Représentation multimodales issue du guide de diagnostic des conduites forcées	150
Figure 48 Fichier Excel exploité comme espace de structuration d'information.....	152
Figure 49 – Mise en forme et analyse de données issues du fichier des indicateurs au travers d'Excel.....	156
Figure 50 Le travail de l'analyste : concevoir des langages .....	158
Figure 51 La saisie procédurale au travers d'un formulaire au regard d'une saisie déclarative dans un document structuré .....	165
Figure 52 L'éditeur de code source souligne en rouge les parties du document non valides	166
Figure 53 L'éditeur de code source, en s'appuyant sur le modèle de document propose les balises valides qui peuvent être introduites .....	166
Figure 54 La capitalisation au travers de documents structurés.....	170
Figure 55 Exemple d'un fichier XML .....	171
Figure 56 Utilisation du document structuré comme support à l'activité de capitalisation des connaissances.....	172
Figure 57 L'éditeur de code source XML Oxygen.....	174
Figure 58 Une vue sur l'interface de saisie proposée par Morphon.....	176
Figure 59 Carte conceptuelle Freemind calculée sur la base d'un fichier XML.....	178
Figure 60 La vision métier du guide de diagnostic dans le tableur OpenOffice .....	179
Figure 61 La saisie sous forme de grille d'un fichier XML dans la version 9 de Oxygen ....	180
Figure 62 Processus de restitution .....	183
Figure 63 Approche de restitution documentaire .....	186
Figure 64 Le guide sous la forme d'un document structuré en XML.....	188
Figure 65 Modèle conceptuel du guide de diagnostic .....	189
Figure 66 Un exemple de graphique à 2 dimensions.....	190
Figure 67 Exemple de visualisation sous forme de graphe.....	190
Figure 68 Du modèle de domaine au modèle de perception .....	192
Figure 69 De l'analyse de l'architexte à la définition de dispositifs .....	196
Figure 70 A gauche la zone de visualisation de documents et à droite la zone de signets ....	197
Figure 71 Exemple de graphe produit par Graphviz.....	198
Figure 72 Exemple de dispositif prédéterminé : le Nœud dans Graphviz .....	200
Figure 73 Exemples de dispositifs personnalisés dans un format pdf : le titre et le tableau.	201
Figure 74 La fenêtre d'application de style dans Microsoft Word.....	202
Figure 75 Du modèle de domaine à l'architexte.....	204

Figure 76 Transformation du document structuré correspondant au guide de diagnostic vers deux visualisations graphiques. ....	205
Figure 77 Mise en relation d'une information avec un dispositif .....	206
Figure 78 Exemple de visualisation graphique du guide de diagnostic.....	206
Figure 79 Sémantique visuelle de construction du document.....	207
Figure 80 Le guide sous la forme d'un fichier PDF adapté à une lecture sur écran et à l'impression .....	208
Figure 81 Le guide sous la forme d'une page HTML .....	209
Figure 82 Vision intégrée de la démarche de gestion des connaissances appuyée par un atelier sémiotique.....	213
Figure 83 Scénario de capitalisation et restitution dans une action de gestion de connaissances.....	215
Figure 84 L'éditeur de code source XML Oxygen.....	217
Figure 85 Une vue sur l'interface de saisie proposée par Morphon.....	218
Figure 86 Editeur d'arbre XML dans Oxygen .....	219
Figure 87 Carte conceptuelle Freemind calculée sur la base d'un fichier XML.....	219
Figure 88 Fichier Excel exploité comme espace de structuration d'information.....	220
Figure 89 La saisie sous forme de grille d'un fichier XML dans la version 9 de Oxygen ....	221
Figure 90 Capitalisation au travers de documents structurés face à un système à base de connaissances existant.....	222
Figure 91 Architecture globale de la maquette implantée .....	224
Figure 92 Exemple de Filtering Table .....	225
Figure 93 Exemple « d'info-bulle » (source : <a href="http://www.walterzorn.com/tooltip/tooltip_e.htm">http://www.walterzorn.com/tooltip/tooltip_e.htm</a> ).....	226
Figure 94 Exemple de galerie photos (source : <a href="http://www.huddletogether.com/projects/lightbox2/">http://www.huddletogether.com/projects/lightbox2/</a> ).....	226
Figure 95 Un éventail de restitutions possibles.....	227
Figure 96 Une capitalisation basée sur une coopération entre experts et analyste.....	229
Figure 97 Une double coopération pour capitaliser et définir la forme de restitution.....	230
Figure 98 La capitalisation étendue aux experts via un site internet de gestion de contenu..	231
Figure 99 Restitutions pour la capitalisation.....	232
Figure 100 Une page du guide de diagnostic dans sa version PDF.....	234
Figure 101 Le guide dans OpenOffice Calc.....	235
Figure 102 Le guide dans OpenOffice Writer.....	236

---

Figure 104 références intermodales.....	237
Figure 104 Exemple de diagramme.....	238
Figure 105 Visualisation sous la forme d'un diagramme accesible au sein d'un navigateur web.....	238
Figure 106 Le document structuré constitue un pivot entre un système existant et un éventail d'architextes.....	244
Figure 107 Capitalisation formelle au travers de documents structurés.....	245
Figure 108 Document structuré en cours de saisie.....	246
Figure 109 Illustration du mécanisme de complétion et de validation en contexte.....	246
Figure 110 Document structuré complété d'un ensemble de balise à remplir.....	247
Figure 111 Traduction du balisage sous la forme de cadres colorés dans l'éditeur Morphon.....	248
Figure 112 une carte conceptuelle produite via Freemind.....	250
Figure 113 Capitalisation hiérarchique de l'information au sein du tableur de la suite bureautique OpenOffice.....	250
Figure 114 Document de capitalisation accessible via le traitement de texte de OpenOffice.....	251
Figure 115 – Association du corpus documentaire et du système à base de connaissances (en tant que donnée d'entrée) pour produire des documents dynamiques interactifs constituant l'interface homme informations.....	252
Figure 116 Processus de génération d'un document multimodal.....	253
Figure 117 Visualisation d'un indicateur technico-économique sous la forme d'un document multimodal.....	254
Figure 118 Document intégrant à la fois une visualisation de graphe et une visualisation tabulaire de la même information.....	254
Figure 119 Une visualisation arborescente et synthétique est intégrée à droite du document textuel pour supporter la navigation.....	255
Figure 120 Document multimodal intégrant 3 formes de représentation : arborescence interactive, graphique et tableau de données.....	256
Figure 121 Visualisation sous forme de graphe du contenu de la base de connaissances.....	257
Figure 122 Visualisation sous forme de graphes de résultats de sortie.....	257
Figure 123 Sémantique visuelle dans la génération de graphe.....	258
Figure 124 Structuration faite par un analyste non informaticien face à une structuration informatique.....	266
Figure 125 L'atelier sémiotique.....	275
Figure 126 Les étapes d'un projet de gestion des connaissances.....	276

Figure 127 A gauche la zone de visualisation de documents et à droite la zone de signets ..	294
Figure 128 La zone des signets .....	295
Figure 129 Exemple de carte conceptuelle produite par Freemind.....	296
Figure 130 Quelques icônes disponibles dans Freemind .....	296
Figure 131 Génération d'un graphe orienté basé sur la décomposition en niveaux d'un graphe .....	298
Figure 132 Génération d'un graphe orienté avec un algorithme de « force direct placement » .....	298
Figure 133 Illustration des clusters dans le graphe et de leur impact sur le traçage d'un graphe orienté.....	299
Figure 134 Exemple de forme de nœuds via Graphviz.....	299
Figure 135 Exemples d'extrémités de liens via Graphviz .....	300
Figure 136 Les tree dans le toolkit Dojo.....	301
Figure 137 Illustration des zones d'édition dans le toolkit Dojo.....	302
Figure 138 Illustration du Titlepane .....	303
Figure 139 – Un document structuré sous-différentes formes .....	306
Figure 140 La saisie procédurale au travers d'un formulaire au regard d'une saisie déclarative dans un document structuré .....	309
Figure 141 Illustration de la plateforme supportant le guide et du réseau d'acteurs dans lequel elle s'insère .....	310
Figure 142 Mécanisme de complétion de balises.....	311
Figure 143 La démarche de formalisation de connaissances avec des documents structurés	312
Figure 144 Le guide sous la forme d'une page HTML .....	313
Figure 145 La personnalisation du guide.....	314
Figure 146 Le guide sous la forme d'un document structuré en XML.....	315
Figure 147 Tissage de liens entre la vision métier et les checks .....	317
Figure 148 Tissage de liens entre la vision métier et l'amdec .....	317
Figure 149 Liens hypertextes créés à partir du tissage de liens entre les fichiers.....	317
Figure 150 Exemple de liens hypertexte et de renvois sur des numéro de page dans une version pdf du guide.....	318
Figure 151 Augmentation en contexte.....	319
Figure 152 Le guide avant augmentation.....	320
Figure 153 Après un click sur « 3-HM... » le détail du check est intégré à la visualisation en cours (cadre rouge).....	320

Figure 154 Illustration du mécanisme d'augmentation basé sur une architecture AJAX.....	321
Figure 155 Exemple de diagramme .....	323
Figure 156 Exemple de diagramme .....	323
Figure 157 Le guide avant intégration d'une zone de mise à jour .....	325
Figure 158 Intégration d'une zone de mise à jour du guide lorsqu'il est consulté.....	325
Figure 159 Intégrer la capitalisation dans le réseau social.....	328
Figure 160 Le guide avant intégration d'une zone de mise à jour .....	329
Figure 161 Intégration d'une zone de mise à jour du guide lorsqu'il est consulté.....	329

---

## Liste des tableaux

Tableau 1 – Niveau d’organisation des variables visuelles .....	121
Tableau 2 Synthèse sur les outils et les ressources produites dans les 2 projets.....	129
Tableau 3 Catégorisation des éléments structuraux et données indentifiables dans l’exemple des guides de diagnostic .....	193
Tableau 4 Synthèse d’analyse du format PDF et SVG .....	199
Tableau 5 Spécification du dispositif « nœud » .....	203
Tableau 6 Définition d’un système de visualisation dans le cas du PDF .....	207
Tableau 7 Définition d’un système de visualisation dans le cas du format HTML .....	208
Tableau 8 Synthèse d’analyse des architextes.....	303

## Partie V - Annexes

## Liste des annexes

## Annexe 1 - Analyse des architextes et de formats de restitution

Cette annexe a pour objet de proposer une analyse plus large d'architextes et de formats de restitution. Il s'agit d'identifier pour chacun d'entre eux d'identifier les moyens de visualisation et d'interaction ainsi que des dispositifs qu'ils proposent. Les exemples traités sont :

- Le format PDF
- Freemind
- Génération de graphe en SVG via GraphViz
- Navigateur WEB et toolkit AJAX

### **Le format PDF et les lecteurs compatibles**

Le format PDF (Portable Document Format) est un format « *de description de page* » (Lux-Pogodalla et Vion-Dury 2004), c'est à dire que « *les entités élémentaires du document sont des caractères, graphiques, images associés à des informations de positionnement sur la page* » (Lux-Pogodalla et Vion-Dury 2004). Avant d'évoquer les variables visuelles propres au format PDF, il est important de mettre l'accent sur le fait que ce format est largement consulté sur écran, même s'il est très adapté à l'impression tout comme le format postscript. Dans un mode de consultation sur écran, le lecteurs exploité va alors implémenter classiquement deux espaces principaux de visualisation du contenu (cf. Figure 127) :

- La zone des signets
- La zone de visualisation du document

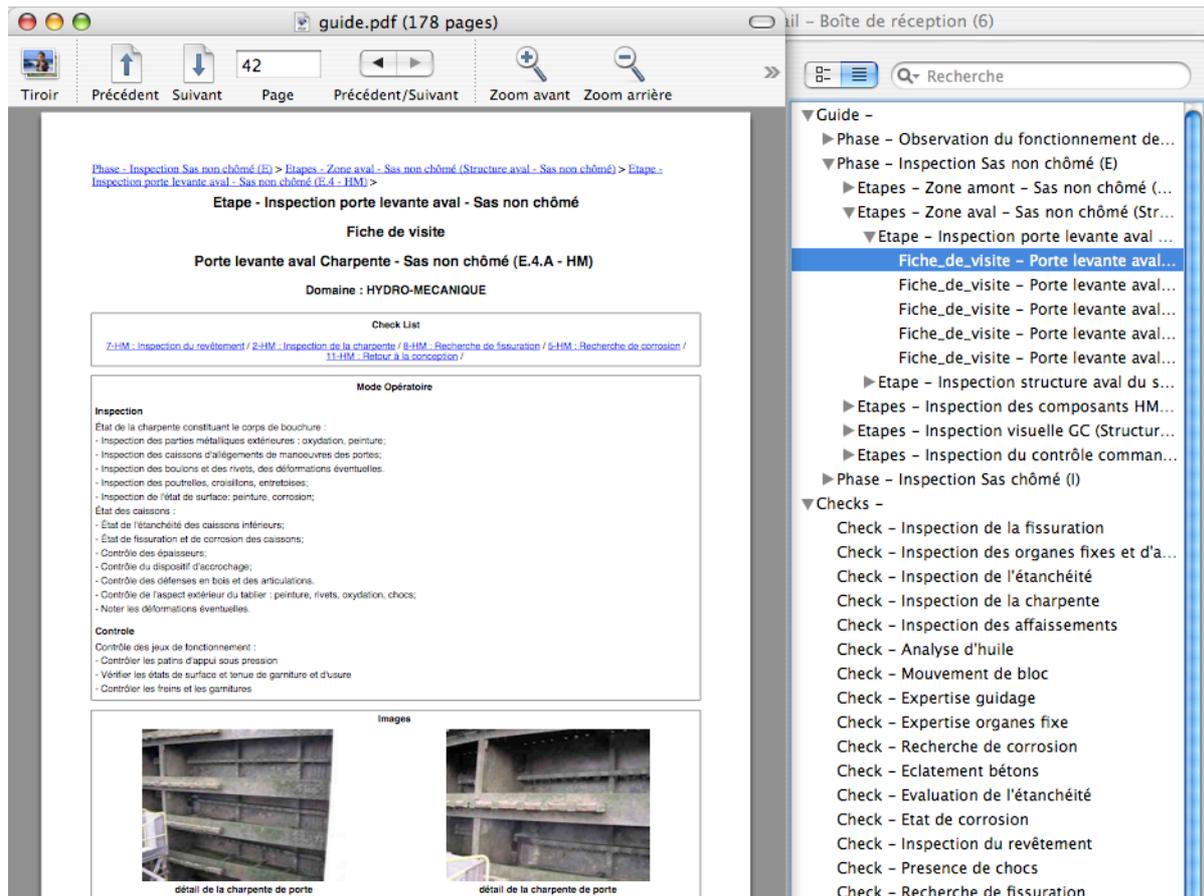


Figure 127 A gauche la zone de visualisation de documents et à droite la zone de signets

La *zone de signets* propose uniquement une visualisation arborescente de texte comme présenté sur la Figure 128. Ainsi, chaque signet correspond à une chaîne de caractère dont il n'est pas possible de paramétrer la mise en forme, puisque celle-ci est en fait laissée à la discrétion de chaque lecteur PDF. La zone de signet propose également deux comportements interactifs ne pouvant être modifiés. Tout d'abord, le signet constitue un lien hypertexte avec la zone de visualisation du document. Ce lien est défini lors de la construction du fichier PDF. Le second mécanisme correspond à l'emboîtement des signets en fonction de l'arborescence. L'emboîtement rend ainsi la visualisation des signets interactive, puisqu'il est possible de rendre visible ou de masquer tout ou partie de l'arborescence.

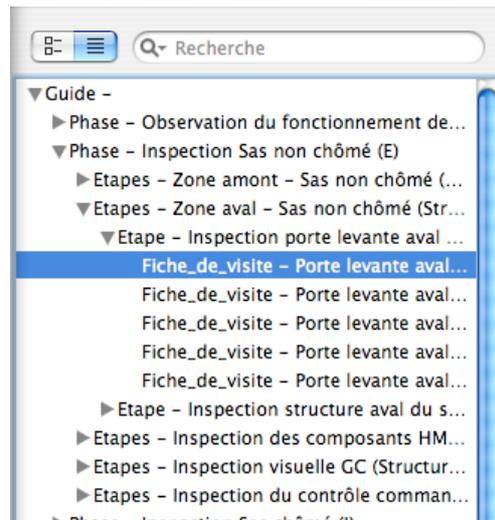


Figure 128 La zone des signets

La zone de visualisation du document offre un espace beaucoup plus créatif. En adéquation avec la définition évoquée précédemment, les variables exploitables sont le texte, avec toute la mise en forme classique (gras, italique, souligner, couleur...), les graphiques et les images. Le plan correspond à la page, et il est possible de positionner les entités qui composent le document assez librement (dans les limites de la page) et de manière très précise : définition de coordonnées claires, saut de page, indentation...

L'interaction dans la zone de texte se traduit tout d'abord par les moyens de navigation dans le document, inhérents au lecteur PDF. Ces moyens sont figés, permettant par exemple le passage d'une page à la suivante, ou encore l'accès direct à une page précise. Au-delà de ces moyens figés, le format PDF permet d'intégrer des liens hypertextes, à la fois intra-documentaire, permettant une navigation non-linéaire dans le document, et inter-documentaire pour mettre en relation ce document avec d'autres entités.

Enfin, et c'est là une possibilité offerte par lecteur PDF de Adobe, dans ces versions les plus récentes, il est possible d'incorporer à un document PDF des annotations. Celles-ci permettent à un utilisateur d'enrichir le contenu du document.

## **Freemind**

Prenons l'exemple de l'éditeur de carte conceptuelle Freemind. Une carte conceptuelle prend la forme d'un diagramme dont la représentation est majoritairement arborescente<sup>7</sup>. Un exemple est présenté sur la Figure 129. Les liens traduisent des connexions sémantiques entre différentes informations. Les informations se traduisent alors par des nœuds contenant du

<sup>7</sup> Ces cartes conceptuelles correspondent plus exactement à des graphes orientés puisqu'il est possible de tracer des liens entre des nœuds contenus dans différentes branches de l'arbre. On retiendra cependant que l'édition de ces liens est parfois difficile lorsque la carte devient grande et ne tient plus alors sur une page écran.



spécifique. On peut également replier l'ensemble des branches d'un nœud d'un simple clic, mais ceci correspond à un comportement imposé.

## **Génération de Graphe en SVG**

Les graphes constituent une forme de représentation classique de l'information relationnelle. Ils permettent d'obtenir une bonne vision de ce type d'information à condition que la manière de les tracer soit correctement implémentée. La bibliothèque Graphviz permet la génération de graphes orientée ou non orientés. La Figure 131 propose un aperçu des résultats qu'elle permet de produire. Les graphes peuvent ainsi être exportés sous la forme d'une image vectorielle en SVG.

Le format SVG est un format d'images vectoriel défini en XML. Les images vectorielles se caractérisent par le fait qu'elles se composent d'une description d'objets géométriques qui permettent de calculer l'image. Elles se différencient des images matricielles (bitmap) qui correspondent à une description de la couleur de chaque pixel composant l'image. Les formats vectoriels d'images ouvrent des possibilités importantes puisque chaque élément qui constitue l'image correspond à un élément identifiable. Un exemple d'exploitation de cette propriété est la mise en relation du texte et de l'image. Dans le cas de la représentation d'un diagramme, il est alors possible de mettre en relation directement les parties du diagramme avec les passages textuels qu'elles explicitent.

Un format vectoriel comme SVG ouvre un champ de possibilité que l'on peut considérer comme illimité en termes de représentation. L'éventail de possibilité n'est par contre réellement exploitable que dans le cadre d'une édition manuelle des représentations graphiques au travers d'un éditeur spécialisé, basé sur la manipulation directe par exemple. Nous évoquerons des éditeurs SVG, comme que Adobe Illustrator ou encore la solution libre Inkscape (<http://www.inkscape.org/>). Dans le cadre d'une génération automatique ou semi-supervisée des représentations graphiques les moyens de visualisation et d'interaction disponibles seront limités par la bibliothèque de génération mise en œuvre.

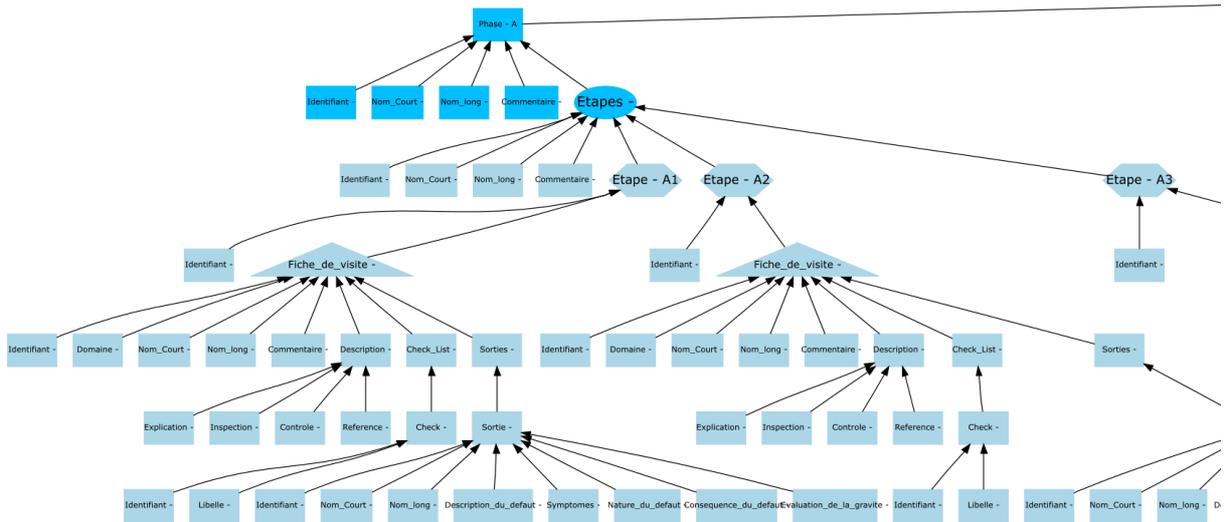


Figure 131 Génération d'un graphe orienté basé sur la décomposition en niveaux d'un graphe

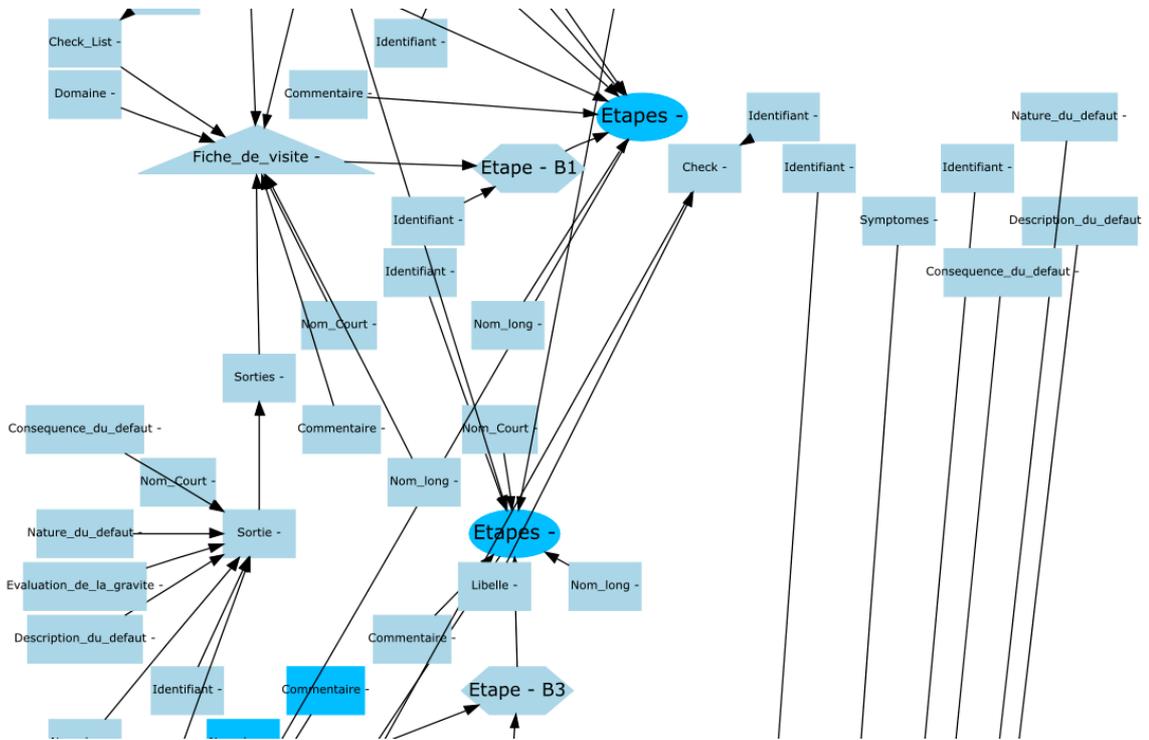


Figure 132 Génération d'un graphe orienté avec un algorithme de « force direct placement ».

Dans le cas de Graphviz, les moyens de visualisation sont principalement en relation avec les nœuds et les arcs du graphe. Il est également possible de spécifier des regroupements de nœuds nommés clusters (cf. Figure 133). Il sera alors possible de faire varier, mais dans une longueur contrainte par l'implémentation de Graphviz, le texte associé, la taille, la couleur ou encore la forme des nœuds et des arcs. La Figure 134 et la Figure 135 présentent un aperçu de l'éventail de formes utilisables.

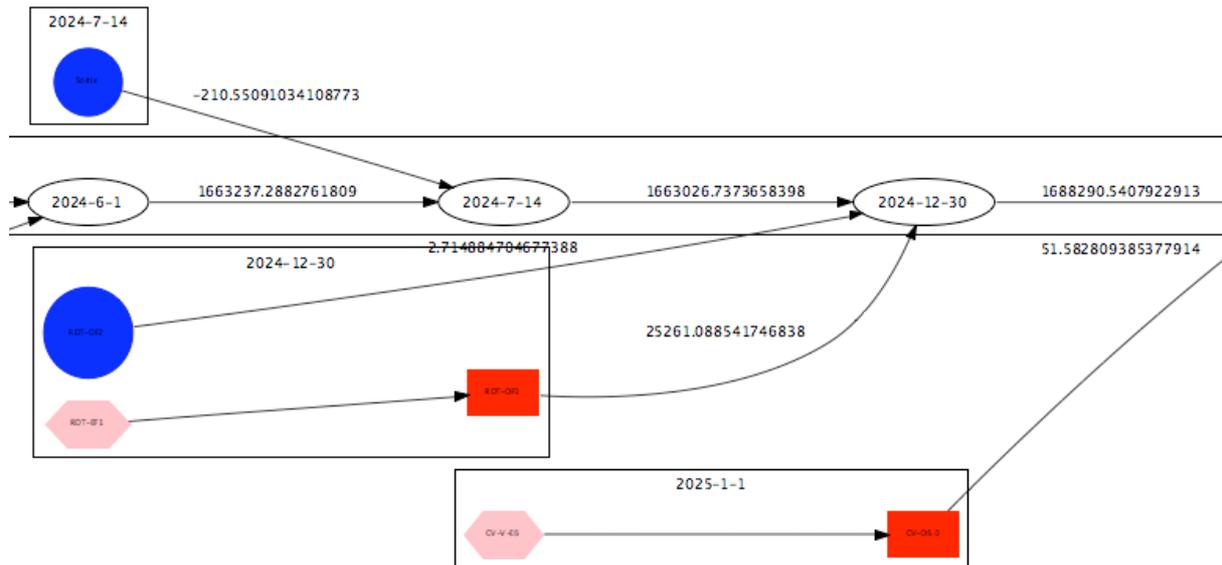


Figure 133 Illustration des clusters dans le graphe et de leur impact sur le traçage d'un graphe orienté

L'utilisation du format SVG permet d'intégrer directement au sein de l'image des mécanismes interactifs tels que le zoom, sans perte de qualité du fait du format vectoriel, et le déplacement dans l'image. La bibliothèque Graphviz apporte pour sa part l'intégration de lien hypertexte lors du calcul de l'image. En complément, nous pouvons évoquer que le format vectoriel ouvre, là encore, un champ de développement quasiment illimité de mécanismes interactifs à associer aux composants de l'image, mais au prix d'un effort de programmation pouvant être important.

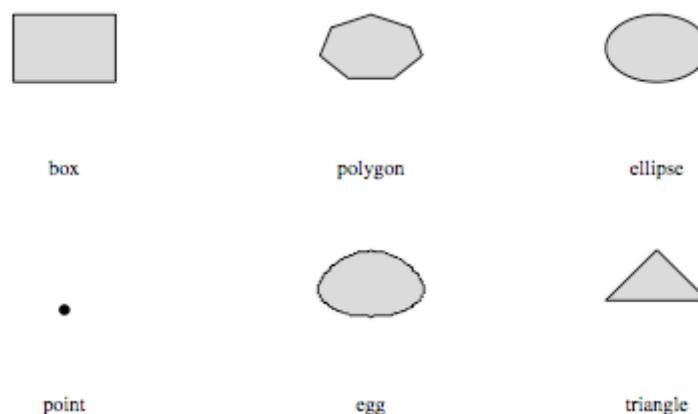


Figure 134 Exemple de forme de nœuds via Graphviz

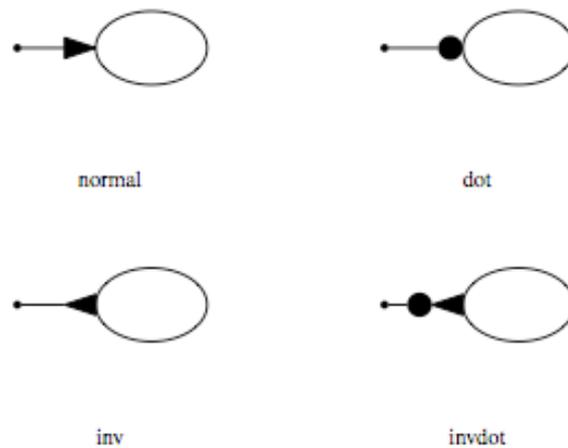


Figure 135 Exemples d'extrémités de liens via Graphviz

### **Navigateur WEB et toolkit AJAX : l'exemple de dojo**

Le navigateur web constitue l'interface de diffusion de contenu la plus répandue. Il est aujourd'hui présent sur l'ensemble des systèmes et des plate-formes existantes. Il s'est armé de plus en plus de fonctionnalités, ne se limitant plus à l'interprétation du code HTML, il intègre pour certains un moteur graphique très puissant permettant l'interprétation de formats tels que le SVG, que nous avons évoqué précédemment, ou encore la lecture directe de son et de vidéo via l'installation de « plug-in ».

Le moteur graphique intégré au navigateur s'attache en premier lieu à la mise en forme du texte et des images. Il permet également de profiter de nombreux moyens interactifs qui se révèlent en fait illimités par le biais de la programmation. Le langage privilégié est le javascript. Ce langage qui n'en est qu'à ses prémices de normalisation et soumis à de nombreuses particularités en fonction du navigateur dans lequel il est interprété. On peut observer le développement de nombreuses bibliothèques aujourd'hui qui facilitent la mise en œuvre de fonctionnalités avancées au sein d'un navigateur souvent compatibles avec divers navigateurs. Ces fonctionnalités qui correspondent aux fondement des Rich Internet Application (RIA), peuvent être appréhendés comme des dispositifs dans notre approche. Nous proposons ici d'analyser 3 dispositifs issus de la bibliothèque javascript Dojo.

- Les tree
- Les zones d'édition
- Les title pane

## Les Tree

Les *tree* sont des dispositifs permettant de représenter une arborescence interactive. Elle se compose donc de nœuds textuels imbriqués hiérarchiquement, dont il est possible de modifier la mise en forme matérielle. Les variations sur chaque nœud peuvent donc se faire sur la taille (ex : taille de police), la valeur, la couleur ainsi que la forme (ex : Police de caractère, gras, italique...). Les moyens d'interaction nativement intégrés à l'arborescence sont les mêmes que ceux évoqués précédemment, c'est à dire les liens hypertextes sur chaque nœud permettant la navigation et l'emboîtement des nœuds en fonction de la hiérarchie, qui permet une modification dynamique de la visualisation (Affichage ou masquage d'une partie de l'arborescence).

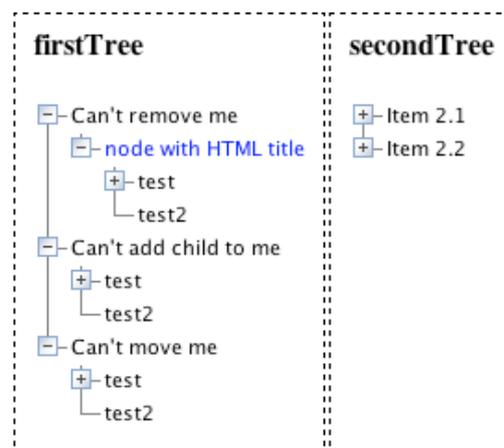


Figure 136 Les tree dans le toolkit Dojo

## Les zones d'édition

L'évolution des navigateurs fait qu'il ne se cantonnent plus à la diffusion d'informations mais proposent des surfaces d'édition permettant d'interagir avec le contenu. Les toolkits proposent ainsi de nombreux dispositifs plus ou moins complexes permettant de transformer une partie d'un document Web en une zone d'édition (cf. Figure 137). Ce type de dispositif en fonction de son niveau de sophistication pourra permettre de mettre en forme le texte, à l'image d'un traitement de texte, permettant de jouer sur de nombreuses variations. Dans leurs versions les plus simples, ils se contenteront d'offrir la possibilité de modifier le contenu d'une balise offrant l'édition comme moyen d'interaction.

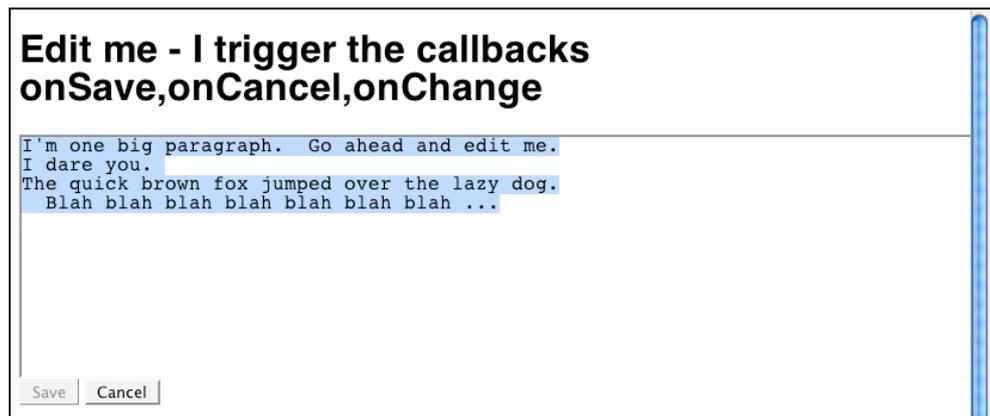
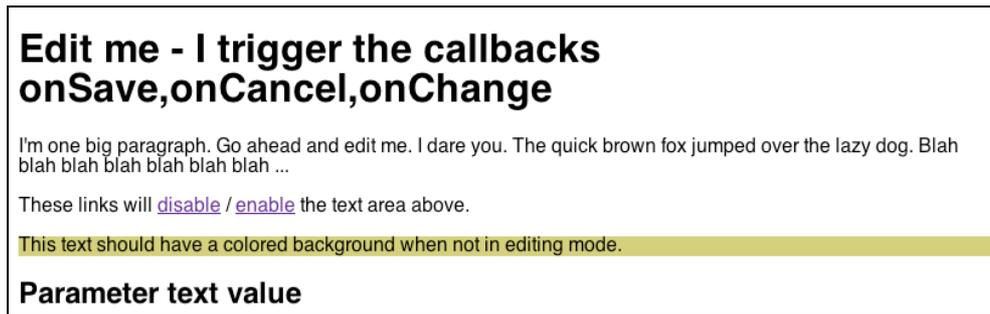


Figure 137 Illustration des zones d'édition dans le toolkit Dojo

### *Les TitlePane*

Le Titlepane est un dispositif composé de deux zones. La première ligne correspondant à ce que nous nommerons le titre, alors que la seconde peut intégrer du texte ou encore de l'image (cf. Figure 138). Au regard du texte ou de l'image, les moyens de visualisation pouvant être exploités sont les mêmes que celles évoqués aux niveaux de la fenêtre du navigateur. Le Titlepane intègre par contre comme moyen d'interaction de pouvoir faire apparaître ou disparaître sa partie contenue, permettant de rendre dynamique la visualisation de l'information.

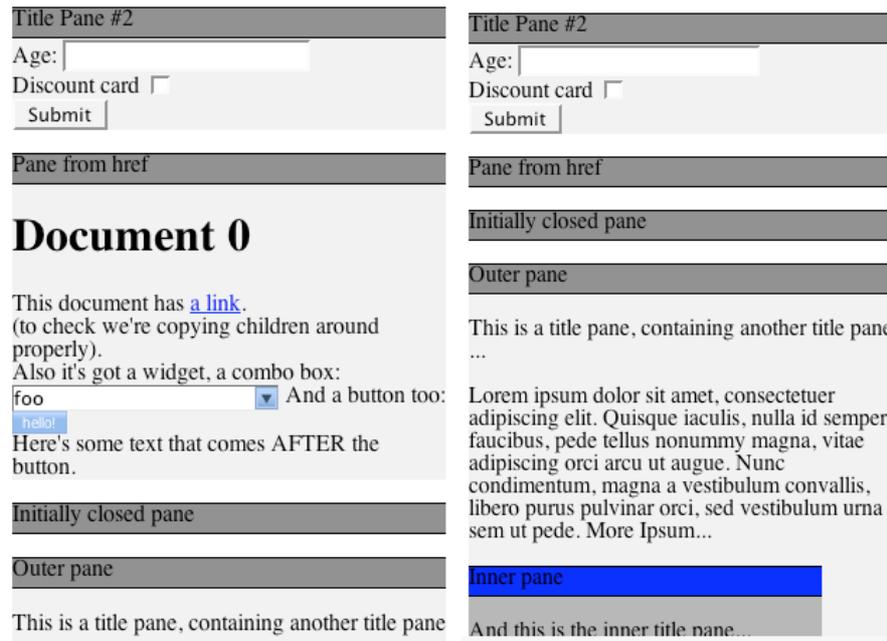


Figure 138 Illustration du Titlepane

Tableau 8 Synthèse d'analyse des architextes

Interface / Format / Architexte	Espace / zone de visualisation et d'interaction	Moyens de visualisation	Variables visuelles exploitables	Moyens d'interaction	Dispositifs
PDF (Acrobat Reader...)	Signets	Visualisation arborescente d'informations textuelles		Liens hypertextes Emboitement des signets	Le Signet
	Document	Visualisation de textes et d'images dans les dimension du plan du document	Plan Taille Valeur Couleur Orientation Forme	Lien hypertexte Eventuellement annotation	A définir
Freemind		Visualisation d'une carte conceptuelle	Taille Valeur Couleur	Lien hypertexte Emboitement des nœuds	Nœuds ou ensemble de noeuds

			Forme	Edition	de la carte
SVG via Graphviz		Visualisation de Graphe dans un plan illimité	Taille Valeur Couleur Forme	Zoom Déplacement (pan) Lien hypertexte Illimité via programmation	Nœud Arc Cluster de Nœuds
Navigateur Web	HTML	Visualisation de textes et d'images dans le plan illimité de la fenêtre du navigateur	Plan Taille Valeur Couleur Orientation Forme	Lien hypertexte Illimité via programmation	A définir
	Toolkit ex : Dojo		Plan Taille Valeur Couleur Orientation Forme	Apparition/Disparition Edition Déplacement Transformation Lien hypertexte ...	Chaque Widget :

## Annexe 2 - Implémentation et mise en œuvre du guide de diagnostic

Cette annexe correspond à un extrait d'un rapport EDF (Grassaud et Parfouru 2007) décrivant le travail effectué dans le cadre du projet relatif aux guide de diagnostic.

### **Le document électronique structuré comme interface de saisie et de restitution du guide de diagnostic**

#### *Définition*

On ne peut réduire le document structuré à la simple évocation de l'utilisation des technologies XML ou plus anciennement SGML. Nous reprendrons les éléments de définition proposés par (Lux-Pogodalla et Vion-Dury 2004) : « *les entités élémentaires sont des éléments structuraux (ou logiques) dont la signification est paramétrable. Les règles d'organisation logique sont explicites et modifiables séparément des instances documentaires (schémas documentaires). La validité structurelle d'un document par rapport à un schéma peut être vérifiée via des algorithmes génériques.* ».

Ces éléments de définition permettent d'appréhender le document structuré comme une structure dont les éléments vont stocker des contenus textuels. Cette structure sera classiquement hiérarchique et sera soumise à une règle d'organisation qui correspond alors à une réelle syntaxe. Ainsi, les différents éléments structuraux, qui définissent les entités élémentaires d'un document associées à leur règle d'organisation, constituent un réel langage. En fonction de la technologie utilisée les mots du langage de structuration sont plus ou moins imposés. Ainsi, HTML (Hypertext Markup Language) définit un langage de structuration de l'information dont les balises sont prédéterminées et répondent à une mise en forme standard. L'intérêt de se tourner vers XML (eXtended Markup Language) est lié au fait qu'il ne constitue pas en soi un langage « *mais plus exactement un ensemble de règles permettant la création de langage de balisage* » (Ray 2001). Ainsi, en s'appuyant sur XML, il est possible d'offrir un espace de structuration de l'information relativement libre dont la seule contrainte est de produire des documents bien formés, c'est à dire en se conformant aux règles du XML.

Le fait de produire des contenus structurés en XML introduit donc deux catégories de documents. Les bien formés que nous venons d'évoquer. Nous avons bien mis l'accent sur l'importance de l'explicitation des règles d'organisation logiques des entités élémentaires et le fait qu'elles doivent pouvoir être validées de manière algorithmique. De là, on peut introduire la notion de document valide. Un fichier XML constituera un document valide si son

organisation structurelle répond à un modèle de document. Ces modèles de document sont typiquement implémentés sous la forme d'une DTD (Document Type Definition) ou d'un Xschema (XML schema). Ils représentent une source d'informations essentielles puisqu'ils définissent un contrat de lecture et de validation des documents, pouvant être appréhendé comme un niveau meta du document.

Un document numérique structuré correspond donc à un fichier balisé dont la structure logique est régie par un modèle de document (Document Type Definition ou Xschema). Ce modèle définit les règles de structuration du document. Sur la Figure 13, nous présentons un exemple trivial de document structuré sous la forme de son codage en eXtended Markup Language (XML). Le choix du format XML induit que l'on se place dans un contexte d'interopérabilité, puisqu'il représente un standard ouvert en dehors de tout formatage propriétaire. Au travers de l'utilisation d'XML, on profite également de l'ensemble des technologies associées à ce standard. La figure présente également ce document comme un arbre ou encore sous une forme sémiotique. Cette dernière correspond à un remplacement du balisage par une mise en forme. Sur l'exemple, les différentes balises ont été remplacées par une coloration.

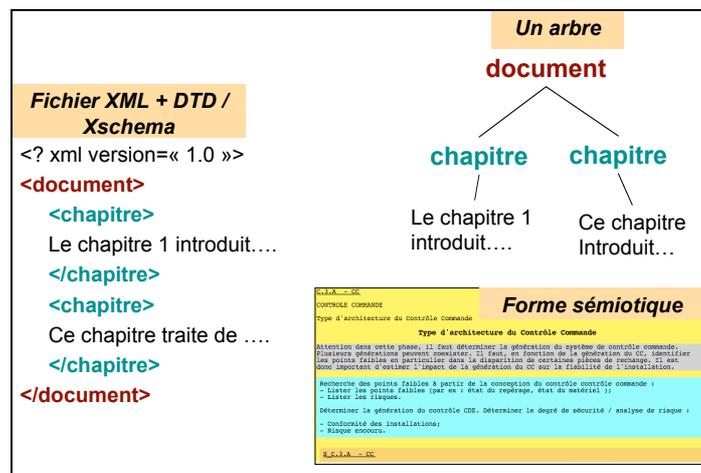


Figure 139 – Un document structuré sous-différentes formes

### *Apports du document électronique structuré comme interface homme machine*

Le choix du document numérique structuré comme interface homme machine est guidé tout d'abord par le fait qu'il introduit une réelle souplesse dans le processus de restitution des données. Il se révèle en effet aisé de passer d'un système de représentation à un autre, voire d'en combiner plusieurs à partir d'un même fichier XML (un document structuré). À partir d'un document codé XML, on peut ainsi produire des documents mis en forme, des schémas, des images...

Au-delà de la forme que l'on souhaite donner au document, la structuration hiérarchique introduite par le codage XML représente une structure de données facilement manipulable. Il est alors possible de filtrer, d'effectuer des recombinaisons, ou encore d'enrichir l'arbre XML que constitue le document.

Lorsqu'on parle d'interface homme information (IHI), il ne s'agit pas seulement de mettre à disposition un contenu, mais également d'être en mesure d'interagir avec ce contenu. Le document numérique structuré constitue une forme de restitution éditable « facilement ». Cette édition peut se faire au travers d'un éditeur de code source XML, qui accompagne la saisie en s'appuyant sur le modèle de document pour en vérifier la cohérence. Il est également possible d'exploiter un éditeur Wysiwyg (« *What you see is what you get* ») qui va occulter les balises, en les traduisant par exemple par un coloriage. Enfin, il est également possible de transformer le document structuré pour qu'il puisse s'adapter à des logiciels standards tels que Word ou Excel. Cette dernière option montre bien l'autonomie du document structuré en tant qu'interface. En effet, il devient indépendant du système avec lequel il a été créé, ce qui introduit des perspectives d'utilisation extrêmement vastes et introduit une certaine pérennité.

*Pourquoi XML dans le processus de formalisation des connaissances : le document structuré et son modèle comme support de structuration de représentation des connaissances*

L'intérêt de l'utilisation de XML en gestion des connaissances n'est pas à démontrer, puisqu'il a déjà été largement discuté, que cela soit en structuration de l'information, ou en restitution de l'information, notamment en valorisant la séparation fondforme permettant de produire à partir d'une même structure différents modes de restitution (Dieng-Kuntz et al. 2001). Nous avons d'ailleurs largement mis à contribution XML dans différents projets (Parfouru et al. 2006; Parfouru et al. 2007).

Pour notre étude, il s'agit de répondre aux inconvénients du détournement de l'informatique bureautique. En effet, nous avons souligné l'importance de l'indexation des informations. Plus cette indexation est fine, et plus on peut imaginer des traitements complexes pour la machine. La production de documents au travers des outils de bureautique, clairement centrés sur le concept de Wysiwyg, produit une information pouvant être structurée par le biais de dispositifs de mise en forme matérielle. La structuration de l'information, ici, se dilue par l'improvisation, introduite par chaque utilisateur dans l'emploi de l'outil : il peut alors y avoir autant de documents que d'individus. Il est envisageable de capter cette indexation des informations mais au prix de post traitements plus ou moins fiables et déterministes. La saisie

des informations directement dans un langage structuré permet de grandement améliorer la situation puisque l'on ne subirait plus une masse documentaire, de plus en plus important, mais on construirait des documents déjà structurés.

De plus, nous nous intéressons plus particulièrement à XML dans l'activité même de structuration des connaissances. En effet, l'activité de capitalisation (renvoyant à la structuration d'une représentation de connaissances) est une activité très peu déterministe. Il est en effet difficile de pouvoir appréhender a priori les informations qui seront effectivement identifiées, de même que leur organisation les unes par rapport aux autres. Le développement de modèles de connaissances, correspondant à des langages de représentation, n'est donc pas une tâche facile, si l'on cherche à développer le modèle en amont.

Notre orientation est d'exploiter le langage structuré XML pour instrumenter l'activité de modélisation. Au travers de l'écriture de documents bien formés, il s'agit de fournir un espace de modélisation qui permette de faire évoluer le modèle de représentation des connaissances et les connaissances capitalisées de concert. L'activité de modélisation correspond à l'écriture d'un document bien formé : le modèle de connaissance correspondant alors au modèle de document en cours d'écriture. Cette démarche doit permettre d'éviter à avoir à anticiper sur le langage de représentation des connaissances, et que ce dernier puisse être construit à mesure du processus de capitalisation.

La tentative est de s'éloigner d'approches qui définissent le modèle a priori et d'interfaces de modélisation trop cadrées et trop contraignantes. On pense notamment aux systèmes de gestion de connaissances dont l'accès est basé sur la mise en œuvre de formulaires. Le formulaire constitue le moyen le plus efficace actuellement « *pour encadrer les auteurs* » à se conformer à un langage formel (Bachimont 2004). « Il propose un guidage méthodologique pour appuyer l'auteur dans sa démarche d'écriture structurée » (Bachimont 2004). Tout ceci est particulièrement pertinent pour gérer la cohérence des informations, mais il constitue surtout une contrainte dans le processus d'écriture. Le guidage méthodologique qu'il introduit, notamment au travers de la succession de formulaires, introduit souvent un ordre précis dans l'écriture des informations. De plus, il est souvent implémenté avec des systèmes fonctionnant sous un mode connecté pour avoir une gestion de la cohérence synchrone avec le système. Au travers du document structuré, nous souhaitons modifier ce mode de saisie qui ne nous paraît pas très propice à l'activité créative que constitue la modélisation. Ainsi, le document structuré peut offrir un espace de saisie tout à fait cadré par un modèle de document, mais dans lequel il sera possible d'écrire les informations selon sa convenance en terme d'ordre par exemple, voire même de laisser de manière temporaire des informations

incomplètes. On profite de l'autonomie du document en introduisant le fait que la gestion de la cohérence peut se faire de manière asynchrone (cf. Figure 51).

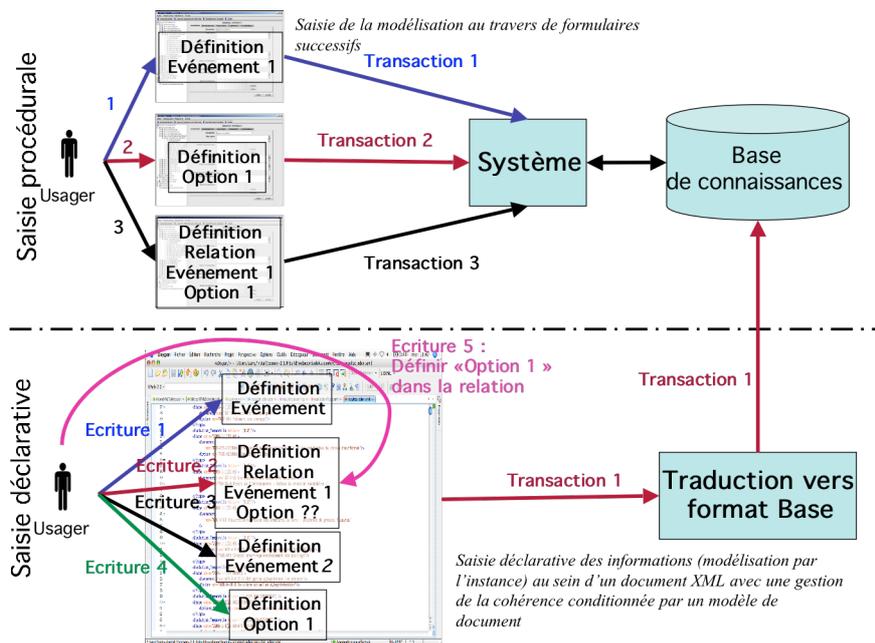


Figure 140 La saisie procédurale au travers d'un formulaire au regard d'une saisie déclarative dans un document structuré

## Implémentation et mise en oeuvre

Nous proposons ici une description de la plate-forme sous-jacente au guide, ainsi que le réseau d'acteurs dans lequel elle s'insère. La Figure 144 propose une représentation des acteurs en précisant les coopérations. Elle traduit, sous la forme de rectangles les différents traitements effectués sur les documents structurés. Enfin, les flèches illustrent les flux de données ou l'interaction entre les acteurs et la plate-forme.

Nous proposons de décrire ici implémentation et mise en oeuvre. On portera tout d'abord un regard en s'intéressant à l'aspect capitalisation des informations. Il s'agit notamment de voir comment les documents structurés sont produits et au travers de quels outils. Par la suite, on abordera l'aspect restitution des documents structurés en décrivant les mécanismes de construction d'une visualisation du guide parfois interactive.

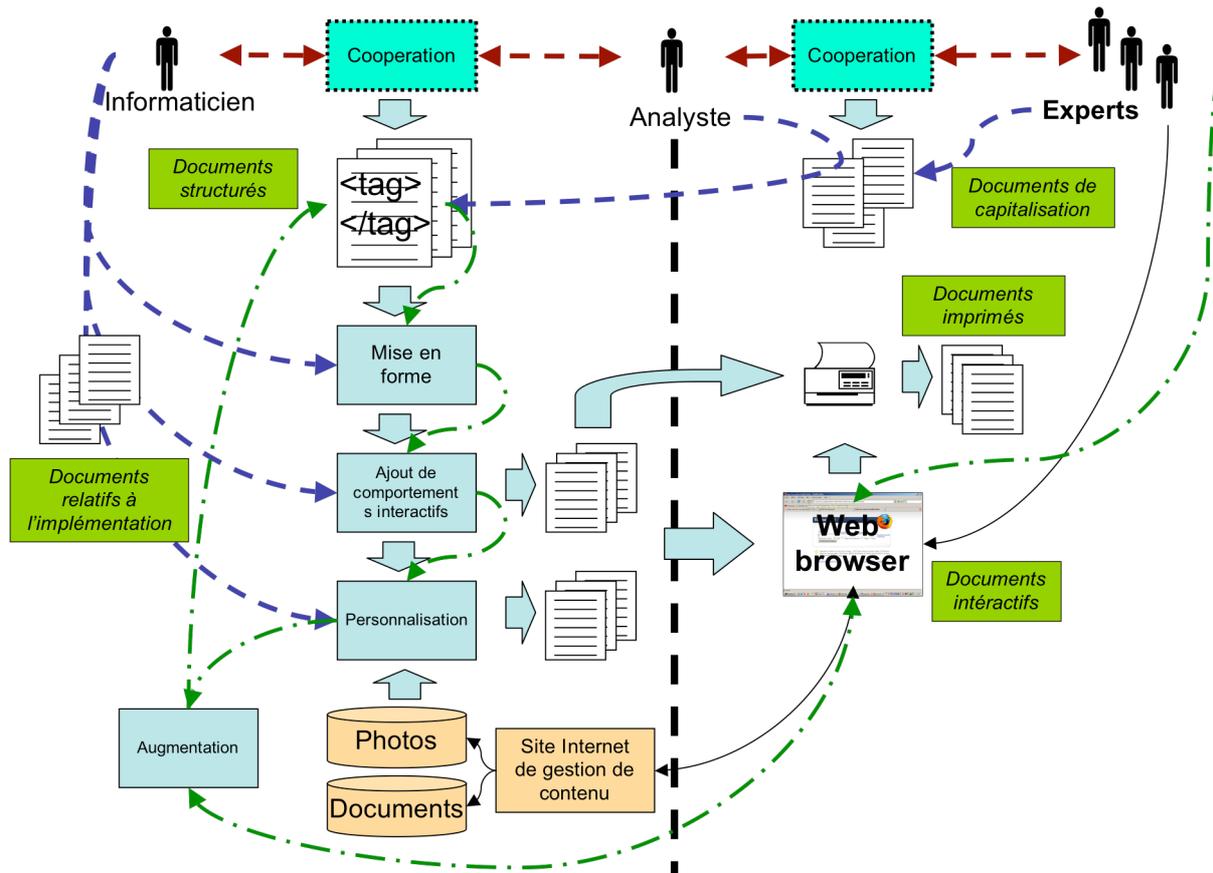


Figure 141 Illustration de la plateforme supportant le guide et du réseau d'acteurs dans lequel elle s'insère

### *Au regard de la capitalisation (démarche expérimentale)*

L'expérience que nous avons mise en place a consisté à mettre un analyste non informaticien, en l'occurrence un sociologue, dans un premier temps face à un éditeur de code source XML. Le choix de l'éditeur s'est porté vers Oxygen (<http://www.oxygenxml.com/>) qui se trouve être un éditeur relativement simple, si on le met en regard d'une référence du marché XMLSpy ([http://www.altova.com/products/xmlspy/xml\\_editor.html](http://www.altova.com/products/xmlspy/xml_editor.html)) distribué par la société Altova.

Cet éditeur offre un espace tout à fait classique d'édition, c'est-à-dire un espace infini d'écriture, ainsi que des capacités de copier-coller relativement importante. Certaines fonctions sont plus liées au langage balisé, puisqu'il introduit des mécanismes de complétion des balises limitant ainsi le travail de saisie des balises. Ce mécanisme de complétion peut se baser, ou sur un modèle de document fourni (DTD / Xschema) ou au travers d'un apprentissage des balises au fur et à mesure qu'elles sont créées dans le document. Il offre également une vérification permanente du caractère bien formé et valide (si un modèle de document est spécifié) du document en cours d'édition.

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <!DOCTYPE Guide SYSTEM "guideEditionEtape.dtd">
3 <Guide>
4 <Phase>
5   <Identifiant>D</Identifiant>
6   <Nom_Court>Observation du fonctionnement de l'ouvrage</Nom_Court>
7   <Nom_long>Observation du fonctionnement de l'ouvrage</Nom_long>
8   <Commentaire>Le moyen principal de surveillance est l'observation visuelle et auditive. Il es
9   Cette phase d'inspection est réalisée par l'exploitant
10  </Commentaire>
11  <Etapes>
12    <Identifiant>Écoute et observation</Identifiant>
13    <Nom_Court>Écoute et observation </Nom_Court>
14    <Nom_long>Écoute et observation</Nom_long>
15    <Commentaire>L'écoute et l'observation de l'ouvrage sans démontage dans les domaines :
16    - Hydromécanique
17    - Genie civil
18    - Controle commande
19
20    </Commentaire>
21
22    <Etape Nom="Observation du fonctionnement hydro-méca ">
23
24      <Identifiant>D.1 - HM</Identifiant>
25      <Fiche_de_visite>
26        <Identifiant>D.1.A - HM</Identifiant>
27
28        <Description>
29          <Texte>
30            <Texte>
31              <Texte>
32                <Texte>
33              </Texte>
34            </Texte>
35          </Texte>
36        </Identifiant>
37        <Texte>
38          <Texte>
39            <Texte>
40              <Texte>
41                <Texte>
42              </Texte>
43            </Texte>
44          </Texte>
45        </Texte>
46      </Fiche_de_visite>
47    </Etape>
48  </Etapes>
49 </Phase>
50 </Guide>

```

Figure 142 Mécanisme de complétion de balises

La prise en main de l'éditeur par l'analyste s'est faite assez rapidement de même que les rudiments du langage XML. La possibilité de définir son propre langage de structuration, permis par l'utilisation de XML, représente une réelle souplesse dans le processus de structuration. L'analyste a ainsi largement apprécié cette manière de structurer ses contenus évoquant même que l'écriture au travers de l'éditeur de code source XML n'était pas très différente de l'écriture dans un traitement de texte. La formalisation des connaissances se fait alors en extension, de manière constructiviste en créant les balises nécessaires au gré des besoins. Structure et contenu évoluent de concert. Le document structuré prend alors une dimension d'espace de conception jusqu'à ce qu'il se stabilise. Cette stabilisation du document permet alors d'identifier le modèle de document – ceci se limite à exploiter les fonctionnalités de l'éditeur, qui permettent de calculer la DTD ou le Xschema du document courant - correspondant à sa structure canonique. Ce modèle de document répertorie alors l'ensemble des éléments structuraux, les balises, qui correspondent au langage de

représentation, de structuration, des connaissances au regard de notre démarche de formalisation (cf. Figure 143).

- Démarche de capitalisation

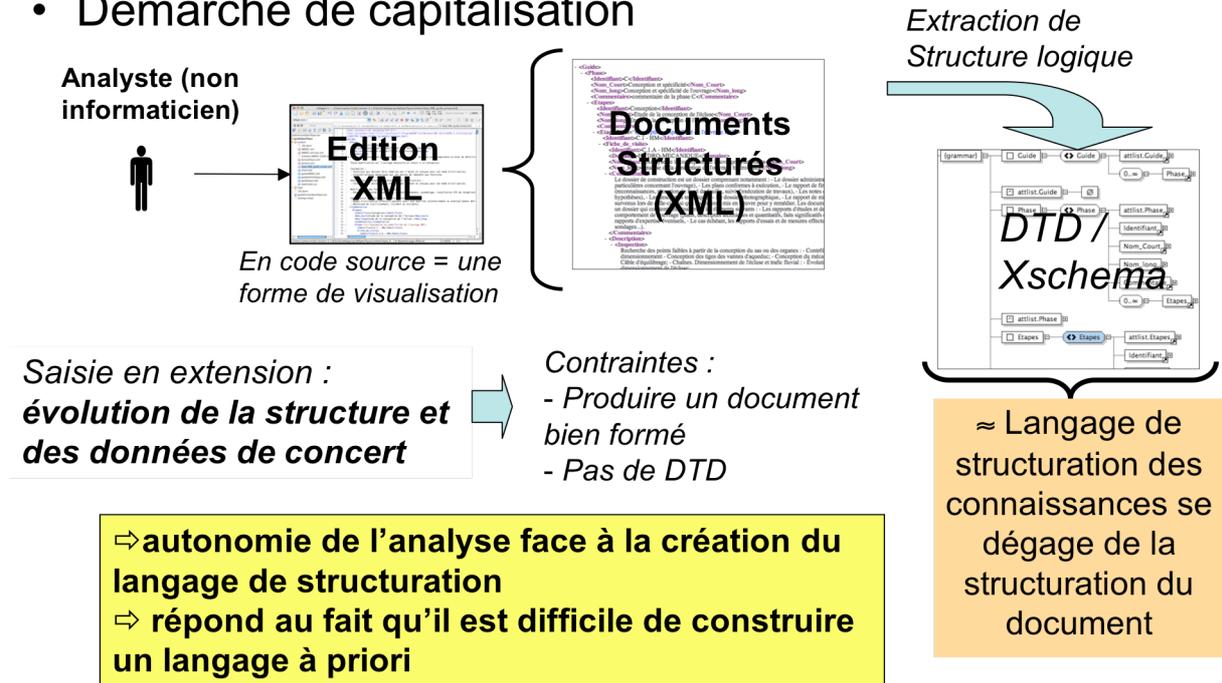


Figure 143 La démarche de formalisation de connaissances avec des documents structurés

Cette démarche et cette instrumentation de la capitalisation n'est toutefois pas sans poser problème. Il est bien évident qu'une formalisation déclarative en extension risque d'introduire des problèmes de cohérence et de « duplication » (répétition) de contenu. C'est en effet un phénomène que nous observons dans nos expérimentations. De même, la structuration hiérarchique introduit nécessairement un point de vue dans l'organisation des informations en rapport avec l'arborescence. On est alors bien loin de la définition d'un modèle conceptuel de données pour aboutir à une base de données permettant l'identification, la localisation et la définition de dépendances tout en palliant la redondance d'informations (pour éviter les incohérences). Tout ceci ne constitue toutefois pas un écueil à nos yeux.

En effet, nous pensons qu'il est important dans le travail d'écriture de l'analyste de produire un tel contenu car il semble difficile de lui imposer de factoriser directement les informations. Par contre, l'intégralité du contenu étant sous une forme très structurée, ces problèmes pourront être traités, a minima de manière semi-supervisé, a posteriori notamment avec l'aide de l'informaticien. Ce dernier sera en mesure d'extraire, avec l'aide de l'analyste, les parties « problématiques » permettant par exemple de créer de nouveaux documents qui factorisent les informations qui se répètent. Ce positionnement est tout à fait lié au fait que nous

exploitons le documents structuré comme un réel espace de conception. Le document produit n'a pas vocation à être directement une version cohérente et optimale. Il constitue un réel instrument de capitalisation et de structuration en produisant une instance . Il peut être appréhendé comme un réel Document pour l'Action (Zacklad 2004), passant d'une version à une autre de part le travail de l'analyste, ou les transformations effectuées par l'informaticien. Ce statut de document pour l'action que prend le document est également lié au fait qu'il constitue un réel objet de coopération, alliant une souplesse de structuration pour répondre aux besoins de l'analyste et une structuration assez importante pour être retravaillée par l'informaticien.

### *Un regard sur la restitution*

Nous proposons d'aborder la restitution des documents structurés produits qui composent le guide de diagnostic. La Figure 144 offre un aperçu du guide, restitué sous la forme d'une page HTML dans un navigateur WEB. Nous allons décrire successivement les mécanismes intégrés au processus de restitution suivant :

- La personnalisation
- Le tissage entre documents
- L'augmentation en contexte

The screenshot shows a web browser interface with a navigation tree on the left and a main content area on the right. The navigation tree includes a 'Guide' section with sub-items like 'Ph-D - Observation du fonctionnement', 'Ph-E - Inspection Sas non chômé', 'Es - Zone\_aval - Inspection porte av', 'Es - Composants\_HM - Inspection c', 'E - Inspection des Bollards - E.7', 'E - Inspection Batardage - E.8 - I', 'E - Inspection Poutre pare-choc', 'E - Inspection organes de vidang', 'E - Inspection vanne aqueduc - E', 'Es - Structure\_passive\_GC - Insp', 'Es - Contrôle\_commande\_CC - Insp', 'Ph-I - Inspection Sas chômé', 'Checks -', 'AMDEC -', 'F - Assurer la navigation -', 'F - Permettre la maintenance à sec de l'c', and 'F - Permettre le franchissement de l'ouv'. The main content area displays a hierarchical structure of colored boxes representing document sections. The top box is blue and titled '- Ph - D - Observation du fonctionnement de l'ouvrage -'. Below it is a green box titled '- Es - Ecoute\_et\_observation - Ecoute et observation -'. Underneath is a pink box titled '- E - Observation du fonctionnement hydro-méca - D.1 - HM -'. Below that is a yellow box titled '- FdV - D.1.A - HM - Fonctionnement hydro-méca de l'écluse -'. Each box contains a 'Commentaire' section with descriptive text. The 'Description' box at the bottom is currently empty.

Figure 144 Le guide sous la forme d'une page HTML

### *Personnalisation*

Une des philosophies de développement du guide d'aide au diagnostic est de diffuser une ressource informationnelle, permettant le partage et la mutualisation des pratiques de maintenance relatives à différents ouvrages. En premier lieu, il doit véhiculer une vision générale du diagnostic. Ceci étant il est également important d'être en mesure de produire une version du guide propre au contexte dans lequel il va être exploité. Nous considérons 3 éléments de contexte pour le moment :

- Le mode d'accès : il s'agit de considérer s'il sera consulté sur écran ou sur papier
- Le domaine de compétence : il contient une vision et une articulation de 3 domaines de compétences. On peut souhaiter se limiter à la consultation de la vision d'un domaine de compétence que cela soit durant la capitalisation ou lors de l'utilisation du guide.
- L'ouvrage : le guide renvoie à une vision générale du diagnostic sur un ensemble d'ouvrages. Chaque ouvrage peut introduire des spécificités notamment concernant leur configuration matérielle (ex : technologies de portes mises en oeuvre)

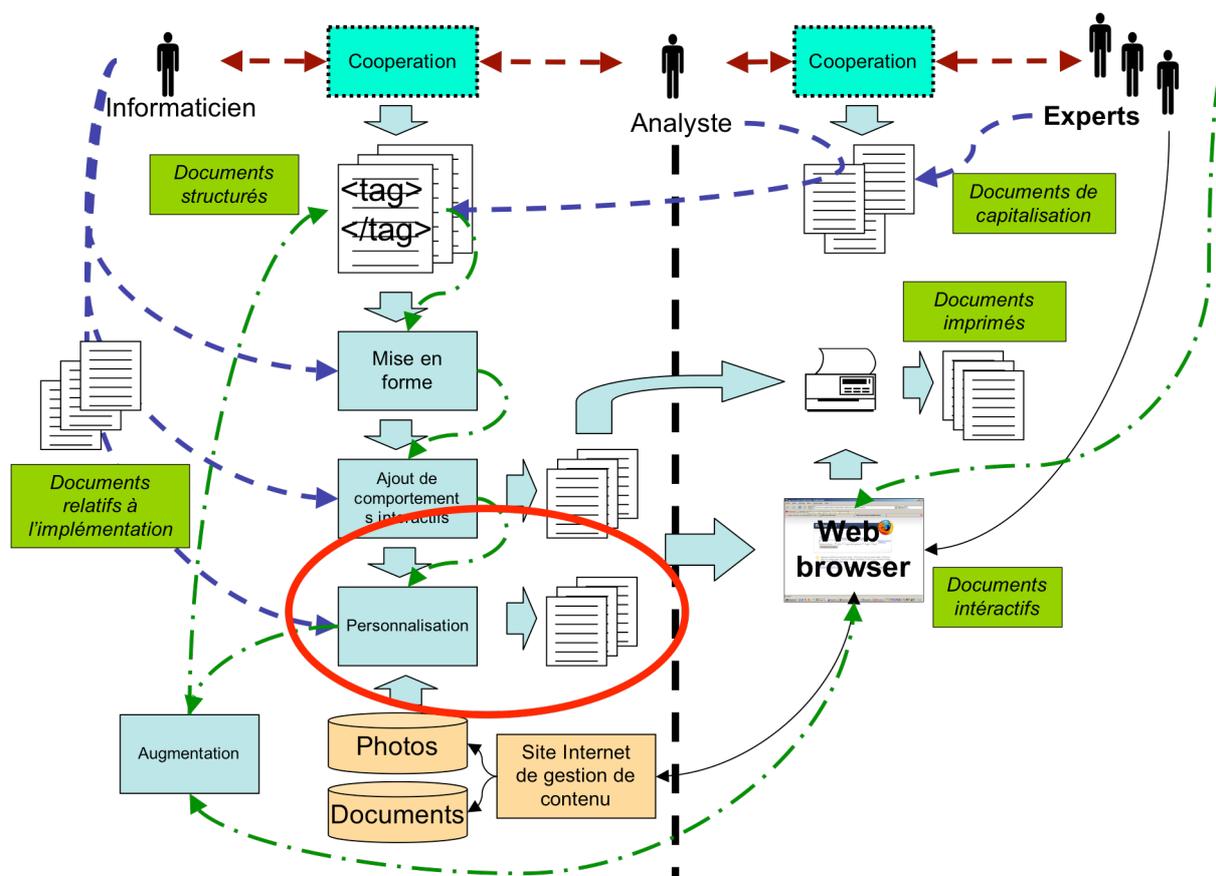


Figure 145 La personnalisation du guide

L'utilisation de documents structurés dans le cadre du guide offre la possibilité, sans engager un effort trop important, de personnaliser le guide sous différentes versions. Ceci est rendu possible de part la définition et l'indexation précise des informations contenues à l'intérieur. Elles sont repérées par des balises qui à la fois sémantisent le contenu et sont adaptées à un traitement informatique. Le document structuré est alors appréhendable comme un arbre, la personnalisation selon le domaine de compétence, ou encore l'ouvrage, revient à élaguer cet arbre. Un aperçu du guide sous la forme d'un document structuré est présenté sur la Figure 146.

L'indentation du texte sur la figure définit la hiérarchie (l'arbre) inhérente au document structuré. Les éléments structurants qui prennent la forme de balises « <element> ... </element> » définissent un découpage du document en fragments identifiables par la machine. Ces fragments peuvent ainsi constituer tantôt des zones de filtrage, tantôt des zones pouvant être élaguées par exemple.

```

- <Guide>
- <Phase Identifiant="C" Nom_long="Conception et spécificité de l'ouvrage" Nom_Court="Conception et spécificité de l'ouvrage">
- <Commentaire>
  Le retour à la conception peut s'avérer indispensable. En effet pour comprendre un état de défaillance, il est parfois nécessaire de remonter à la
  conception du composant mais aussi à sa technologie, des matériaux utilisés ou du dimensionnement initial. Toute modification sur l'ouvrage nécessite
  un retour à la conception. Conception : - Fonctions qui doivent être remplies par l'objet et conçues pour son mode d'utilisation. - Caractéristiques
  mécaniques qui ont permis de répondre aux fonctions. Réalisation et réception (éventuellement modification) : - Fonctions qui doivent être remplies
  par l'objet et conçues pour son mode d'utilisation. Modifications des fonctions conçues : - Réception en terme de composant, approvisionnement,
  assemblage, installation (PV de réception). Exploitation et utilisation (usure et vieillissement) : - Modes d'utilisation effectifs pouvant avoir été
  modifiés volontairement ou éventuellement dérive des réglages. - Mécanisme de vieillissement, incident et accidents.
</Commentaire>
- <Etapes Identifiant="Conception" Nom_long="Etude de la conception de l'écluse " Nom_Court="Etude de la conception de l'écluse ">
  <Commentaire>
- <Etape Identifiant="C.1-HM" Nom_long="Conception et spécificité de l'ouvrage HM">
- <Fiche_de_visite Identifiant="C.1.A-HM" Domaine="HYDRO-MECANIQUE" Nom_long="Analyse de la conception de l'ouvrage
  hydro-méca" Nom_Court="Analyse de la conception de l'ouvrage hydro-méca">
  HYDRO-MECANIQUE
- <Commentaire>
  Le dossier de construction est un dossier comprenant notamment : - Le dossier administratif ( concession et avenant, cahier des charges
  particulières concernant l'ouvrage), - Les plans conformes à exécution, - Le rapport de fin de chantier, - Les rapports géologiques
  (reconnaisances, réception de fond de fouille, suivi d'exécution de travaux), - Les notes de calcul (pour permettre de connaître les
  hypothèses), - Les descriptifs techniques, - Le dossier photographique, - Le rapport de mise en eau et en particulier, les incidents éventuels
  survenus lors de celle-ci ainsi que les moyens mis en oeuvre pour y remédier. Les documents de la vie de l'ouvrage, ces documents constituent
  un dossier qui contient principalement les éléments suivants : - Les rapports d'études et de travaux d'entretien, de modification et de
  comportement de l'ouvrage (plans, descriptifs techniques et quantitatifs, faits significatifs des études et du chantier), - Les fiches de visites et les
  rapports d'expertise éventuels, - Le cas échéant, les rapports d'essais et de mesures effectués {hydrauliques, caractéristiques des matériaux,
  sondages...). - Le cahier des charges de la concession ou un extrait de la partie relative au canal. - Le règlement général de la police de
  navigation. - Le code des voies navigables et de la navigation intérieure. - Le code de navigation propre à la voie d'eau considérée. - La liste
  des jours légaux de chômage de la navigation. - Fiche ouvrage et dossier ouvrage
</Commentaire>
- <Description>
- <Inspection>
  Recherche des points faibles à partir de la conception du sas ou des organes : - Contrôler le poids de la vanne et refaire les calculs de
  dimensionnement - Conception des tiges des vannes d'aqueduc; - Conception du mécanisme de manoeuvre de la porte aval et amont; -
  Câble d'équilibrage; - Chaînes. Dimensionnement de l'écluse et trafic fluvial : - Évolution du dimensionnement des bateaux; - Défaut de

```

Figure 146 Le guide sous la forme d'un document structuré en XML

### *Tissage entre documents*

Dans sa matérialisation, le guide est en fait composé de trois documents structurés différents:

- La vision métier,
- Les checks,

- L'AMDEC.

L'originalité et l'apport de cette version électronique du guide est de mettre en relation les différents documents structurés, ce qui permet de croiser différents points de vue. On peut par exemple la vision métier avec l'AMDEC qui renvoie à un point de vue construite selon le découpage fonctionnel de l'ouvrage. Cette mise en relation est envisageable au travers d'une rédaction manuelle du document et a d'ailleurs demandé un effort important sans compter qu'il peut être difficile de rendre le tout cohérent.

Le guide intègre un mécanisme algorithmique de tissage de liens entre les fichiers. Des liens sont alors créés entre la vision métier et les checks (Figure 147) et l'AMDEC (Figure 148). Il en résulte une structure pouvant être aisément manipulée et valorisée par de nombreux traitements informatiques. Ce tissage permet la création d'un hypertexte : on intègre ainsi au sein de la vision métier des liens hypertextes vers l'AMDEC et inversement. Cette solution est utilisée aussi bien lorsque l'on génère le guide sous la forme d'un document WEB (Figure 149) que sous la forme d'un fichier PDF, les liens étant utilisables lors de la consultation sur le support numérique. Pour ne pas perdre cette dimension interaction lors du passage au papier, les documents produits, en particulier dans le format PDF, font également mention au niveau de référence de la page vers laquelle il renvoie (cf. Figure 150). Ceci permet de conserver la lecture non linéaire dont document en dehors du monde numérique.

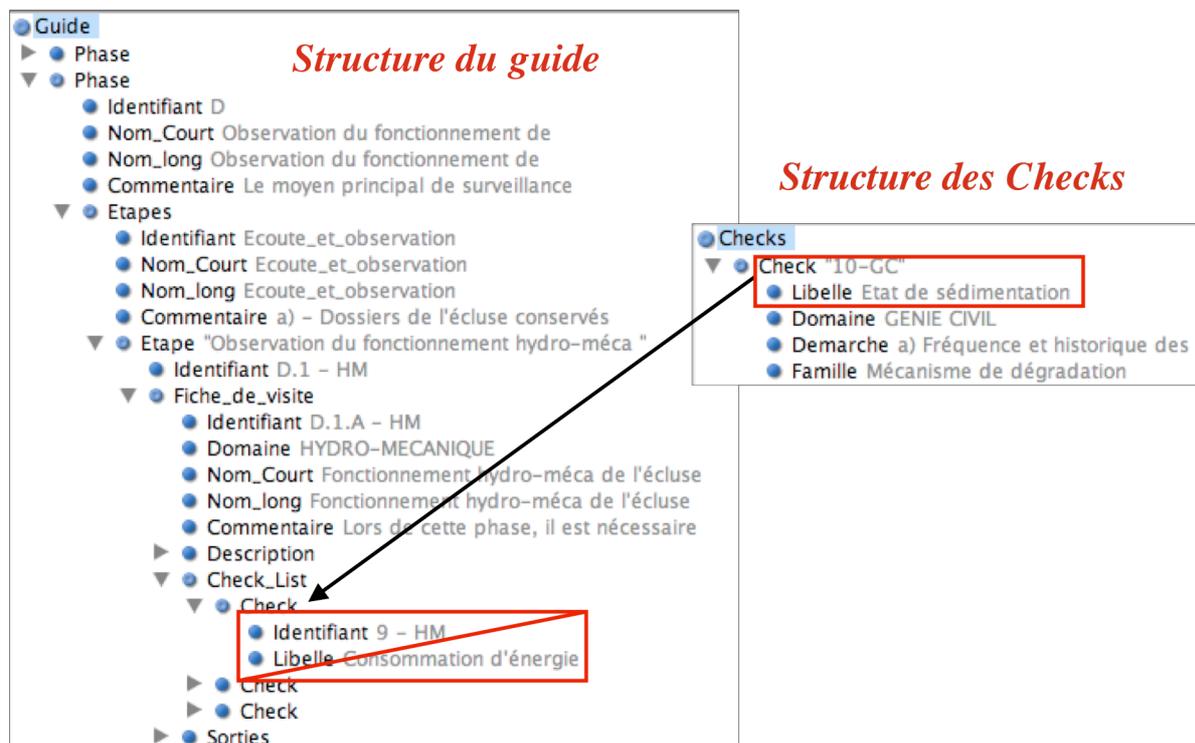


Figure 147 Tissage de liens entre la vision métier et les checks

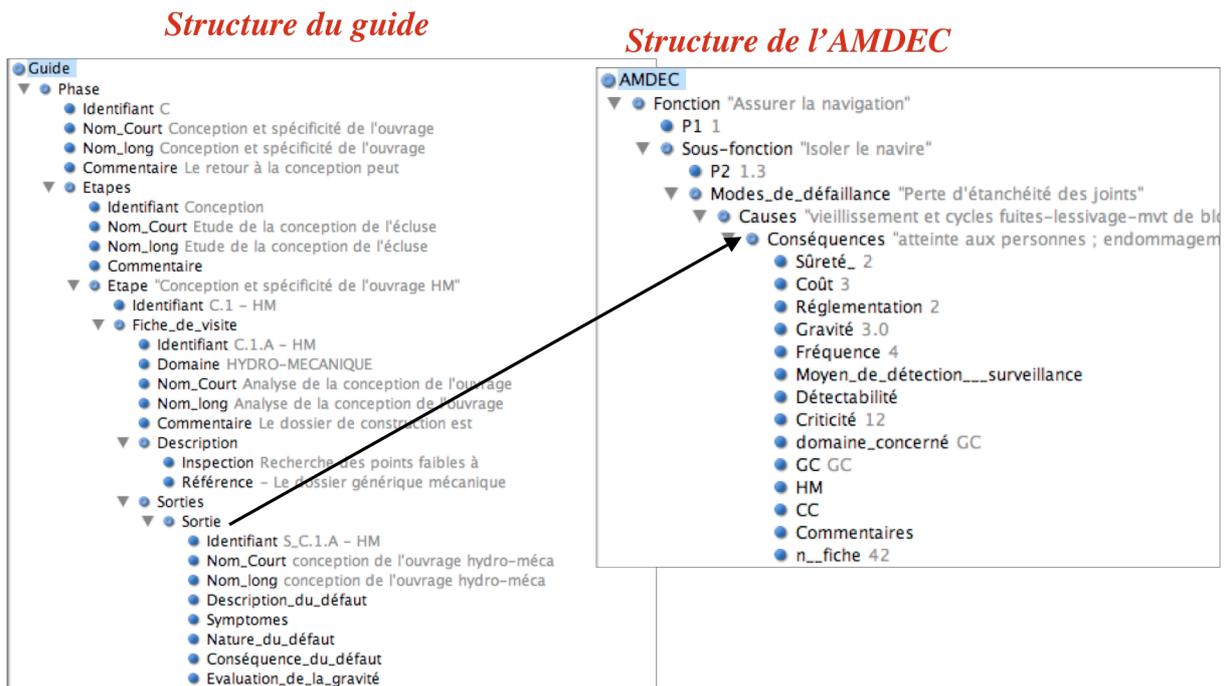


Figure 148 Tissage de liens entre la vision métier et l'amdec

**MD - fuites dans le sas**

Causes - perte d'étanchéité

Composant :

Conséquences - surfaces glissantes sur le navire (gel) et mise en danger navigants ( N° fiche : 81)

Indicateurs : S : - Co : 1 - R : 2 - G : 3.0 - F : 4 - Cr : 12

Détection :

[sas non chômé - Bajoyer de sas](#) >> [S\\_E.1.A-GC](#)

[sas non chômé - Mur de chute](#) >> [S\\_E.3.A-GC](#)

[sas non chômé - Ouvrages de drainage, puits d'accès et remblais](#) >> [S\\_E.4.A-GC](#)

[sas chômé - Fondations](#) >> [S\\_I.1.A-GC](#)

[sas chômé - Radier](#) >> [S\\_I.2.A-GC](#)

[sas chômé - Vannes aqueduc](#) >> [S\\_I.6.A-GC](#)

[sas chômé - Bajoyers de sas](#) >> [S\\_I.7.A-GC](#)

[sas chômé - Mur de chute](#) >> [S\\_I.8.A-GC](#)

[sas chômé - Puits de contre poids](#) >> [S\\_I.9.A-GC](#)

**Dans l'AMDEC**

**MD - chutes d'objets**

- Auscultation des écluses de Fessenheim / Bilan 2003-2004 et adaptation - I/H/FESS-6 ERS/PS04/00001/A BPE.

Constat : Evaluation du risque de défaillance - Ouvrages de drainage, puits d'accès et remblais (S E.4.A-GC)

**Dans le guide**

Mode de défaillance associées

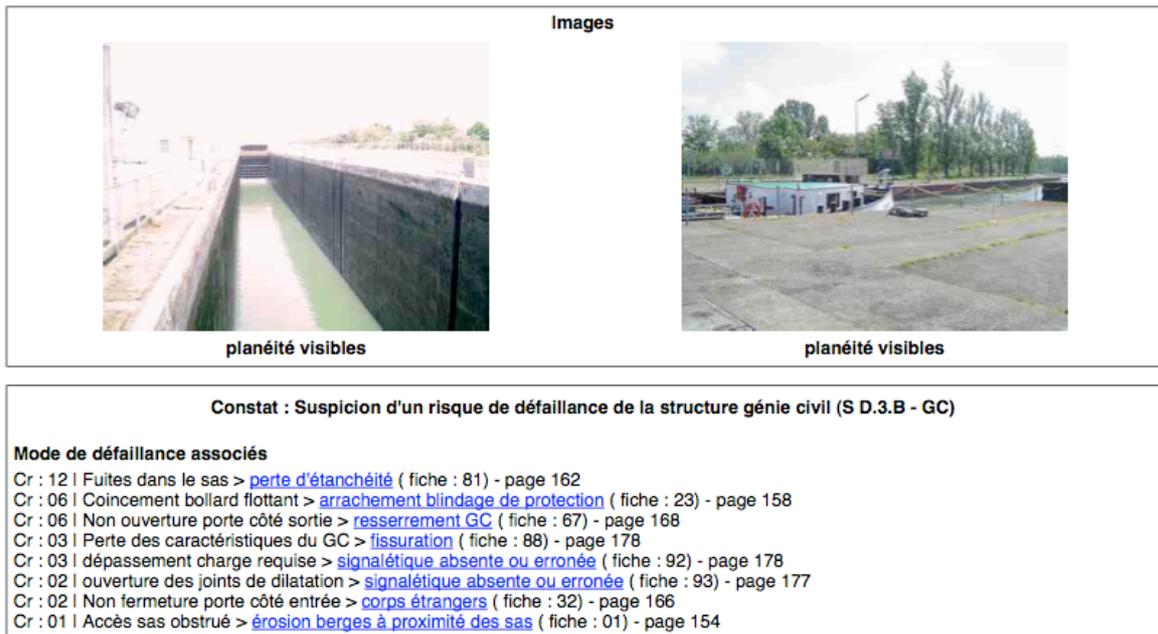
Cr : 12 | [fuites dans le sas](#) > [perte d'étanchéité](#) > [81](#)

59

Cr : 9 | [dégradation de la structure GC](#) > [cycle perte d'étanchéité -lessivage du sol de fondation - perte de capacité portante](#) > [41](#)

Cr : 6 | [perte de capacité portante au niveau des remblais](#) > [drains bouchés ou effondrés](#) > [111](#)

Figure 149 Liens hypertextes créés à partir du tissage de liens entre les fichiers



11

Figure 150 Exemple de liens hypertexte et de renvois sur des numéro de page dans une version pdf du guide

### *Augmentation en contexte*

Le tissage de liens que nous venons d'évoquer peut être valorisé sous une autre forme que des liens hypertextes. La navigation dans un hypertexte introduit un mode de lecture de l'information basé sur la navigation. L'inconvénient est que le passage d'un écran à un autre lors de l'invocation d'un lien – on nomme cela le changement de contexte – peut générer des phénomènes de désorientation dans l'espace informationnel. Nous proposons alors la mise en œuvre d'un mécanisme d'*augmentation en information* en contexte.

Le mécanisme d'augmentation en information permet lorsqu'on consulte une partie du guide d'y intégrer de manière dynamique et interactive les parties de guide qui étaient auparavant pointée par les liens. Ceci est illustré par les Figure 152 et Figure 153. La première présente une vue du guide consultée dans un navigateur internet. La seconde présente la même vue sur laquelle on a invoqué le mécanisme d'augmentation en information : sur cette seconde vue, on identifie une zone cadrée, qui présente le détail d'un « check » sur lequel on a cliqué.

Le mécanisme d'augmentation en information constitue un mode d'accès particulier au guide qui permet, d'une certaine manière, de se construire son propre document. En effet, un utilisateur pourra vouloir ou non faire apparaître les détails relatifs aux checks ou à l'AMDEC

référéncés dans la vision métier et ainsi se construire sa propre visualisation, sa propre version du document.

L'implantation du mécanisme se traduit de la manière suivante (illustré par la Figure 154). Nous avons mis en place une architecture Ajax (Asynchronous JavaScript And XML) permettant, à partir d'un document chargé en mémoire dans le navigateur, d'importer de manière asynchrone dans ce document des informations issues d'autres sources. Le mécanisme est défini avec une stratégie rudimentaire, puisqu'il s'appuie en fait uniquement sur les identifiants pour localiser les informations connexes à afficher. Dans une version plus évoluée, il serait possible d'introduire un moteur de recherche permettant par exemple d'importer des informations issues de tout un corpus documentaire.

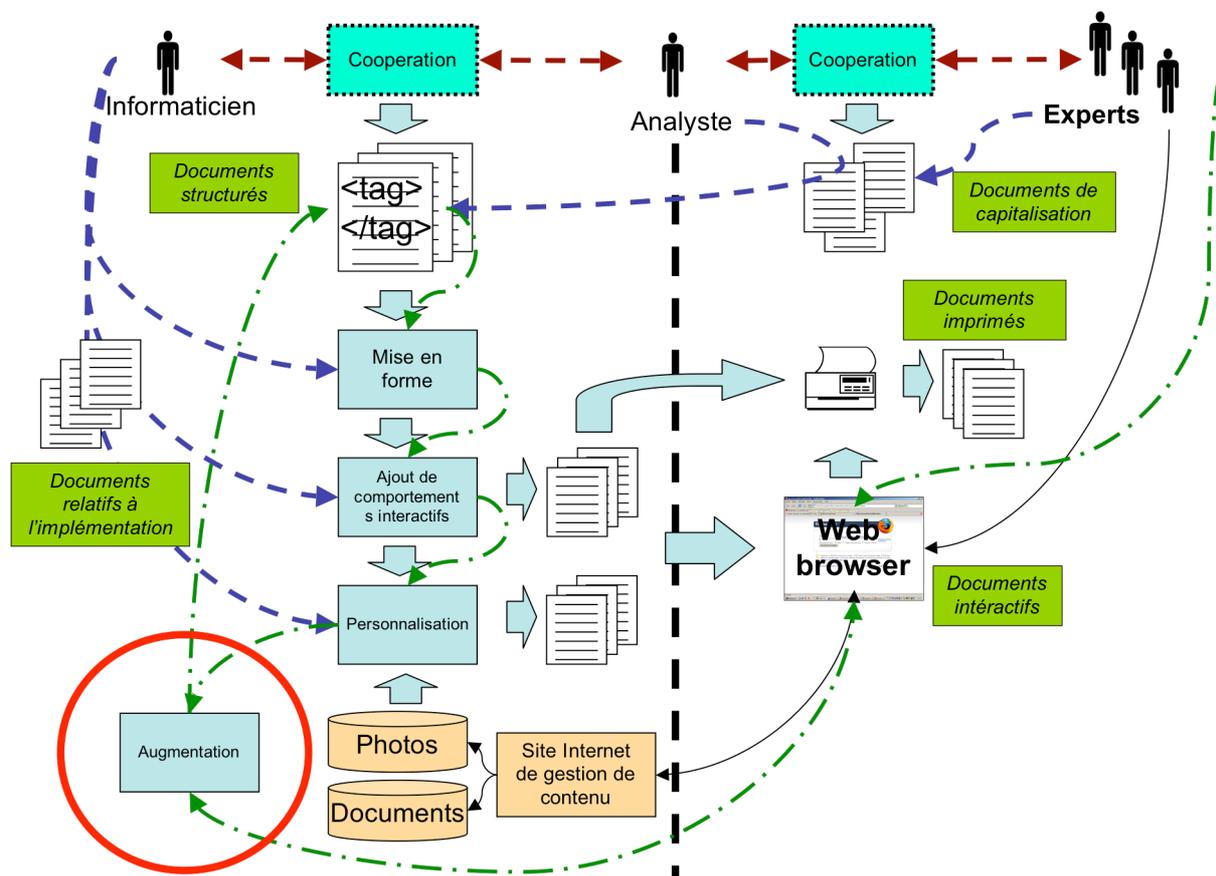


Figure 151 Augmentation en contexte

**- Guide -**

- Ph - D - Observation du fonctionneme
- Ph - E - Inspection Sas non chômé -
- Es - Zone\_ament - Inspection porte
- Es - Zone\_aval - Inspection porte av
- Es - Composants\_HM - Inspection c
- E - Inspection des Bollards - E.7
- E - Inspection Batar dage - E.8 - t
- E - Inspection Poutre pare-choc
- E - Inspection organes de vidang
- E - Inspection vanne aqueduc - E
- Es - Structure\_passive\_GC - Inspect
- Es - Contrôle\_commande\_CC - Insp
- Ph - I - Inspection Sas chômé -
- Checks -
- AMDEC - -
- F - Assurer la navigation - -
- F - Permettre la maintenance à sec de l'e
- F - Permettre le franchissement de l'ouv

**Porte levante** **Bollard**

**- Check List -**

9-HM - Consommation d'énergie 3-HM - Analyse d'huile 7-HM - Inspection du revêtement

**- S - S\_D.1.A - HM - Suspicion d'un risque de défaillance hydro-mécanique -**

**- Modes de défaillances associés -**

- Cr : 9 | efforts résistants trop importants ( fiche : 65 )
- Cr : 9 | corps étrangers au niveau des roulements ( fiche : 25 )
- Cr : 9 | choc navire ou poids lourd sur porte ( fiche : 39 )
- Cr : 9 | choc navire ou poids lourd sur porte ( fiche : 51 )
- Cr : 9 | coincement galets ( fiche : 24 )
- Cr : 9 | corps étrangers au niveau des roulements ( fiche : 25 )
- Cr : 9 | pompes défailantes ( fiche : 104 )
- Cr : 9 | rupture partie d'ouvrage ( fiche : 80 )
- Cr : 9 | efforts résistants trop importants ( fiche : 36 )

Figure 152 Le guide avant augmentation

**- Guide -**

- Ph - D - Observation du fonctionneme
- Ph - E - Inspection Sas non chômé -
- Es - Zone\_ament - Inspection porte
- Es - Zone\_aval - Inspection porte av
- Es - Composants\_HM - Inspection c
- E - Inspection des Bollards - E.7
- E - Inspection Batar dage - E.8 - t
- E - Inspection Poutre pare-choc
- E - Inspection organes de vidang
- E - Inspection vanne aqueduc - E
- Es - Structure\_passive\_GC - Inspect
- Es - Contrôle\_commande\_CC - Insp
- Ph - I - Inspection Sas chômé -
- Checks -
- AMDEC - -
- F - Assurer la navigation - -
- F - Permettre la maintenance à sec de l'e
- F - Permettre le franchissement de l'ouv

**Porte levante** **Bollard**

**- Check List -**

9-HM - Consommation d'énergie 3-HM - Analyse d'huile

**- Check - 3-HM - Analyse d'huile -**

**Domaine :** HYDRO-MECANIQUE

**Demarche :**

- Recueil de la date de la dernière vidange d'huile.
- marque et référence de l'huile utilisée
- Prélèvements et analyses d'huile des organes suivants :
  - Réducteurs, ponts, renvois d'angle;
  - Valeur de la viscosité de l'huile, dosage de l'eau, présence de particule de matière en suspension, couleur de l'huile et ancienneté de l'huile, présence de particules de bronze;
  - Caractéristiques mécaniques de l'huile;
  - Aptitude de l'huile à remplir sa fonction;
  - Recueil de la date de la dernière vidange d'huile ainsi que l'huile utilisée.
- Analyses physico-chimiques
- Analyses spectrométriques des huiles

**Famille :** Mécanisme de dégradation

7-HM - Inspection du revêtement

**- S - S\_D.1.A - HM - Suspicion d'un risque de défaillance hydro-mécanique -**

Figure 153 Après un click sur « 3-HM... » le détail du check est intégré à la visualisation en cours (cadre rouge)

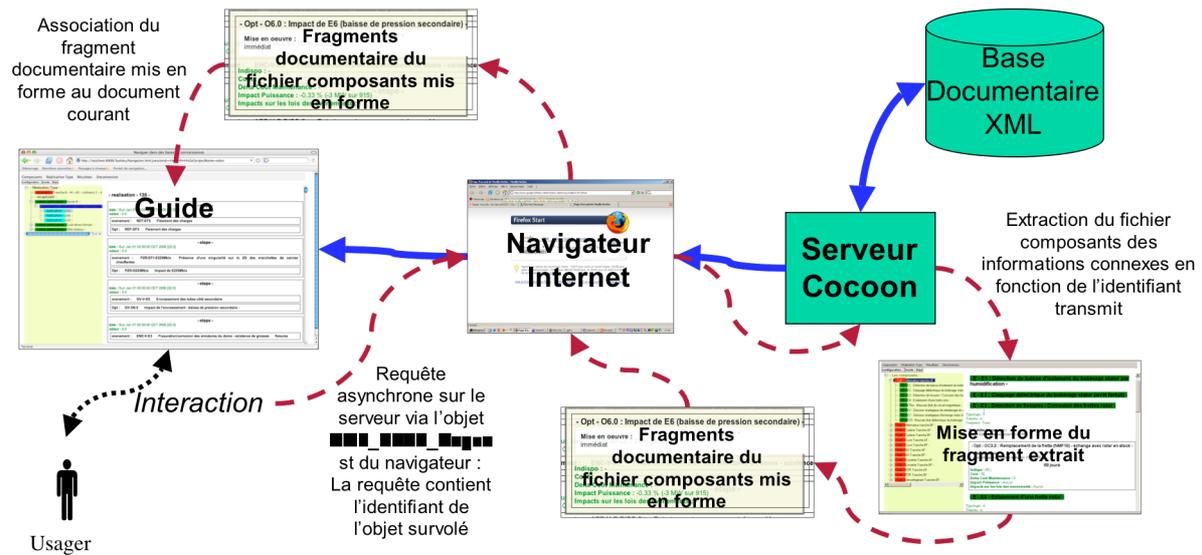


Figure 154 Illustration du mécanisme d'augmentation basé sur une architecture AJAX

## Évolutions et perspectives

En prolongement de ce qui vient d'être décrit, nous pouvons proposer un certain nombre d'évolutions et perspectives liées au développement du guide et à son intégration dans les activités métiers. Certains des éléments que nous venons d'évoquer ont fait l'objet d'expérimentations que nous n'avons pu approfondir plus en amont ou font l'objet d'expériences en cours.

### *Au regard de la restitution*

La constitution du guide sous la forme d'un document structuré permet, à partir d'une unique source de données, d'en produire différentes versions au regard de la forme et du fond. Nous évoquons trois pistes d'approfondissement :

- L'intégration de Diagrammes
- La réorganisation du contenu en fonction des criticités<sup>8</sup> de l'AMDEC
- La réorganisation du contenu en fonction d'une action métier

L'*intégration de diagrammes* a pour vocation de compléter le contenu du guide qui est aujourd'hui majoritairement sous une forme textuelle. Les diagrammes avaient déjà été exploités dans les premières versions des guides de diagnostic produits au travers de Microsoft Word et Microsoft Visio. Le diagramme est une représentation graphique qui doit permettre un accès différent à l'information, comparé au texte, qui est consulté dans la

<sup>8</sup> La notion de criticité correspond à un facteur d'évaluation du risque

continuité du document, majoritairement de manière linéaire et séquentielle. Le diagramme doit également permettre de donner une vision globale du diagnostic qui est aujourd'hui relativement absente du guide. Il doit ainsi se positionner comme une cartographie facilitant la navigation et la compréhension du contenu du guide. Une démarche pour produire des diagrammes a déjà été initiée, permettant d'aboutir à des représentations telles que présentées sur la Figure 155 et la Figure 156.

Nous proposons également deux types de réorganisation. La première que nous évoquons correspond au fait de réorganiser la vision métier du guide en fonction des criticités de l'AMDEC par exemple. Le tissage de liens entre le point de vue métier et le point de vue fonctionnel correspondant à l'AMDEC est intéressant pour la navigation hypertextuelle ainsi que pour le mécanisme d'augmentation en information que nous avons vu précédemment. En complément, le fait d'avoir à disposition les criticités de l'AMDEC au niveau des fiches pourrait permettre par exemple d'ordonner le contenu du guide en fonction des criticités. On pourrait ainsi caractériser certaines parties du diagnostic (donc certaines fiches) comme plus ou moins cruciales. Cette indexation fine et l'intégration des criticités pourrait d'ailleurs venir appuyer un système d'aide à la décision ou une gestion de maintenance assistée par ordinateur (GMAO).

Un second type de réorganisation est de paramétrer l'accès au contenu du guide en fonction d'une tâche métier particulière. Prenons l'exemple du rondier qui va scruter, analyser l'ouvrage, en fonctionnement ou non. Il parcourt l'ouvrage selon un chemin plus ou moins défini. Il est imaginable de permettre au rondier de définir le parcours qu'il va effectuer afin de produire une version du guide en relation avec celui-ci. Ainsi, cette version spécifique, produite dynamiquement, pourrait être le résultat d'un filtrage du guide afin de ne faire apparaître que les fiches pertinentes dans son parcours. Il s'agirait également de réordonner les fiches de visite sélectionnées conformément au parcours défini. Le guide ne serait plus seulement une source d'informations constituant un référentiel, mais également un réel instrument support à l'activité métier.



Figure 155 Exemple de diagramme

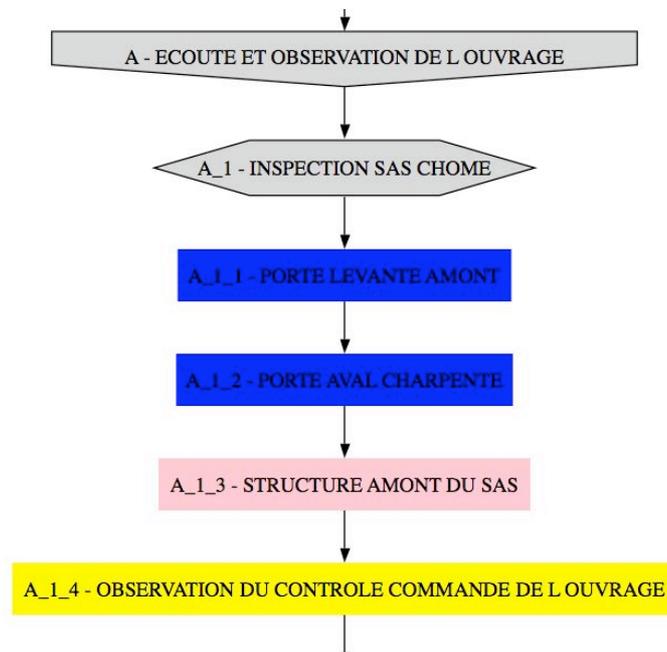


Figure 156 Exemple de diagramme

*Au regard de l'interaction*

La consultation du guide par le biais d'une interface informatique et non pas sous une forme papier permet de proposer des mécanismes interactifs à l'utilisateur. Nous avons évoqué

précédemment l'utilisation de liens hypertextes et la mise en œuvre d'un mécanisme d'augmentation en information. Nous proposons ici d'en décrire deux autres qui pourraient être intégrées au guide :

- La navigation dans le guide au travers de représentations de l'ouvrage
- La mise à jour en ligne du guide

L'organisation du contenu du guide est essentiellement basée sur la modélisation du diagnostic. Cette modélisation le structure sous la forme d'une hiérarchie composée de phase, d'étapes et de fiches de visite qui influencent largement la navigation dans le guide. Il convient de la comprendre pour aisément parcourir son contenu et trouver l'information dont on a besoin. Ce mode d'accès, qui peut être appréhendé comme logique reste toutefois centrée sur des personnes ayant une bonne connaissance de la structuration du diagnostic. Nous proposons une méthode alternative, s'appuyant sur le balisage du contenu, qui permet en effet d'affiner les recherche en spécifiant les zones jugées comme pertinentes. Il serait ainsi possible de naviguer dans le guide selon de nombreux critères ou stratégies et en particulier les parties (portes, bati ...) qui composent l'ouvrage. On pourrait alors proposer un plan de l'ouvrage qui constituerait alors une carte de navigation dans le contenu. Le fait de cliquer sur une partie de l'ouvrage offrirait alors l'ensemble des éléments de diagnostic (les fiches de visite) correspondant à la partie pointée.

Si offrir des mécanismes de navigation constitue un élément important, la mise à jour et le maintien du contenu l'est pas moins. Force est de constater que la mise à jour du contenu est encore faite au niveau R&D actuellement par les concepteurs du guide eux-même. La problématique est qu'il semble difficile de proposer aux experts de modifier les documents structurés sous leur forme brut en code source. Pour répondre à ce problème, il est envisageable de fournir une mise à jour en ligne directement lors de la consultation du guide au niveau du navigateur web. Les Figure 157 et Figure 158 illustrent la transformation d'une zone du guide directement en zone de saisie et ce, de manière interactive.

**E - Observation du fonctionnement hydro-méca - D.1 - HM -**

**- FdV - D.1.A - HM - Fonctionnement hydro-méca de l'écluse -**  
 Domaine : HYDRO-MECANIQUE

**Commentaire :**  
 L'écoute du fonctionnement sans démontage :  
 - Placé à proximité des organes en mouvement (système de manoeuvre des portes, des vannes aqueduc, accouplement, pallier), il s'agit d'écouter et observer les composants en fonctionnement comme par exemple la chaîne cinématique de manoeuvre des organes etc.

**- Description -**

**Explication :**  
 Ecoute des phénomènes acoustiques anormaux :  
 - Écoute des vibrations (réducteurs, renvois d'angle, accouplements) , des bruits de grincement, des cliquetis, des frottements, des bruits de roulements etc. qui peuvent être le signe de désordres mécaniques.  
 - En situation normale, le fonctionnement des organes doit être fluide.  
 - Tout bruit anormal peut être le signe de dents casées sur les pignons, chaînes battantes, mauvais réglages ou jeu anormal, Noter les vibrations et les bruits des engrenages pendant le fonctionnement :  
 - Vibration des ventaux (problème de jeux ou mauvaise répartition des forces);  
 - Analyse vibratoire en cas de suspicion.

**Inspection :**  
 Portes :  
 - État du fonctionnement et du réglage des portes;  
 - Temps de manoeuvre des portes;  
 Mécanisme de manoeuvre des portes :  
 - Détection de fuite d'huile;  
 - Contrôler l'efficacité du graissage des crapaudines;  
 - Contrôler le niveau d'eau dans les puits de contrepoids;

Figure 157 Le guide avant intégration d'une zone de mise à jour

**E - Observation du fonctionnement hydro-méca - D.1 - HM -**

**- FdV - D.1.A - HM - Fonctionnement hydro-méca de l'écluse -**  
 Domaine : HYDRO-MECANIQUE

**Commentaire :**  
 L'écoute du fonctionnement sans démontage :  
 - Placé à proximité des organes en mouvement (système de manoeuvre des portes, des vannes aqueduc, accouplement, pallier), il s'agit d'écouter et observer les composants en fonctionnement comme par exemple la chaîne cinématique de manoeuvre des organes etc.

**- Description -**

```
<content class="text">Ecoute des phénomènes acoustiques anormaux : </content>
<content class="text">- Écoute des vibrations (réducteurs, renvois d'angle, accouplements) , des bruits de grincement, des cliquetis, des frottements, des bruits de roulements etc. qui peuvent être le signe de désordres mécaniques. </content>
<content class="text">- En situation normale, le fonctionnement des organes doit être fluide. </content>
<content class="text">- Tout bruit anormal peut être le signe de dents casées sur les pignons, chaînes battantes, mauvais réglages ou jeu anormal, Noter les vibrations et les bruits des engrenages pendant le fonctionnement : </content>
<content class="text">- Vibration des ventaux (problème de jeux ou mauvaise répartition des forces); </content>
<content class="text">- Analyse vibratoire en cas de suspicion. </content>
```

Save Cancel

**Inspection :**  
 Portes :  
 - État du fonctionnement et du réglage des portes;  
 - Temps de manoeuvre des portes;

Figure 158 Intégration d'une zone de mise à jour du guide lorsqu'il est consulté

### *Adaptation à des outils existants*

Nous venons d'aborder la problématique d'édition de contenu en proposant une solution s'appuyant sur l'exploitation d'un navigateur internet. Le navigateur apporte un avantage au niveau du déploiement des applications, puisqu'il s'agit d'une application quasiment disponible sur l'ensemble des plate-formes auxquelles nous sommes confrontés. Ceci représente donc un bon vecteur pour la diffusion du guide et son utilisation. De plus, le

navigateur s'insère, pour le moment, dans une architecture client-serveur. Le caractère centralisé de ce déploiement facilite la mise à jour, que cela soit au niveau du contenu, mais également au niveau des fonctionnalités proposées. L'utilisateur du guide accède à une version toujours à jour de ce guide dès lors qu'il n'en a pas enregistré une copie localement.

Cette situation qui peut paraître convenable pose toutefois le problème d'avoir accès au réseau pour exploiter le guide en ligne. On parle souvent d'informatique ubiquitaire, mais force est de constater qu'un accès permanent au réseau, d'une entreprise par exemple, reste difficile à imaginer et à garantir d'autant que le coût peut en être très important. Une approche alternative est de fournir le guide sous une forme qui soit autoportante, c'est à dire qui ne nécessite pas un serveur central pour fonctionner. L'idée est alors de produire un guide qui soit compatible avec les logiciels d'une suite bureautique. La plus communément déployée reste aujourd'hui la suite Microsoft. Le caractère propriétaire de cette suite logicielle impose un format qui n'est pour le moment pas ouvert, ce qui rend difficile de s'y conformer. Aussi, dans un premier temps nous proposons d'expérimenter la possibilité d'accéder au guide via la suite OpenOffice et en particulier son traitement de texte ainsi que son tableur. L'ambition est alors de produire une version du guide qui sera modifiable, annotable via les mécanismes de révisions. Ces informations seront alors exploitables pour assister la mise à jour du guide à partir du terrain.

### *Liens avec des ressources informationnelles existantes*

Le guide doit représenter une aide au diagnostic sous la forme d'une vision relativement « synthétique ». Son objectif est de fournir une description, ni trop précise et ni trop exhaustive. Il s'agit plutôt de préciser une démarche. Dans son contenu, le guide introduit parfois des références bibliographiques, en rapport avec les éléments de diagnostic concernés. Il est alors possible d'imaginer que le guide se positionne comme un moyen de navigation dans de nombreuses ressources informationnelles existantes. Il s'agirait de le mettre en relation avec des bases documentaires, permettant lors de sa consultation d'avoir accès, par exemple, à des documents qui y sont référencés (par exemple les MECEP). Dans le même ordre d'idée, de nombreux résultats d'études relatifs aux ouvrages sous la forme de données de calcul pourraient également être mis à disposition et contextualisés. Il s'agirait également de d'intégrer ou de lier les plans d'ouvrages numérisés disponibles. Il en va de même de l'intégration de photos, qui permettent de fournir une information qui n'est pas transmissible par le texte : le texte décrit alors que la photo montre. Le guide pourrait alors devenir un

carrefour à partir duquel on accéderait à de nombreuses ressources numériques et dans lesquelles il serait facile d'être désorienté.

### *Intégrer la capitalisation dans le réseau social*

La démarche de capitalisation des connaissances actuellement mise en place est majoritairement centrée sur l'analyste. Il mène les entretiens et se charge de l'ensemble de la saisie au sein du guide. Il semble que ce mode de fonctionnement était tout à fait pertinent pour initier le travail et arrivé à une version du guide présentant une certaine stabilité, il s'agit ensuite d'intégrer la ressource développée dans les activités métier. En particulier, le guide doit représenter une ressource vivante, au sens où son contenu doit pérenniser une forme de représentation, mais également pouvoir évoluer. Cela semble ne pouvoir se faire de manière efficace que si les personnels de terrain, en particulier les experts, ont la possibilité d'enrichir son contenu.

Nous initiions ainsi une expérience en mettant à disposition un site web collaboratif sur lequel les experts ont la possibilité de déposer des photos et des documents relatifs aux différentes écluses. L'objectif est de partager les informations disponibles et de les indexer. Par exemple, pour une photo, on demande de renseigner sur quel ouvrage elle a été prise, ainsi que le ou les composants visibles sur celle-ci. Nous nous plaçons ici dans une indexation sociale. Cela signifie que ce n'est plus l'analyste qui va construire et saisir l'ensemble des informations mais bien des personnels de terrain : on est alors dans une logique de partage et de mutualisation de l'information directement au niveau terrain.

A partir d'une telle ressource, nous pouvons développer des algorithmes pouvant enrichir ou personnaliser les versions du guide générées avec documents et photos indexés sur le site collaboratif. Une illustration peut être que lorsque l'on génère le guide spécifique à l'ouvrage de Strasbourg, l'ensemble des photos contenues dans le guide soient effectivement des photos relatives à Strasbourg.

Le site de partage de contenus présente un autre intérêt important. Il peut sauvegarder une réelle mémoire de l'évolution de l'ouvrage. Par exemple, le dépôt des photos est à chaque fois horodaté. Dans le guide, il est alors envisageable de faire apparaître une suite de photos relatives à un même composant d'un ouvrage particulier pour analyser ou prendre connaissance de son évolution (ex : l'évolution d'une fissure dans le génie civil). Ce type

d'approche pourrait permettre d'instrumenter de manière intéressante l'activité de maintenance.

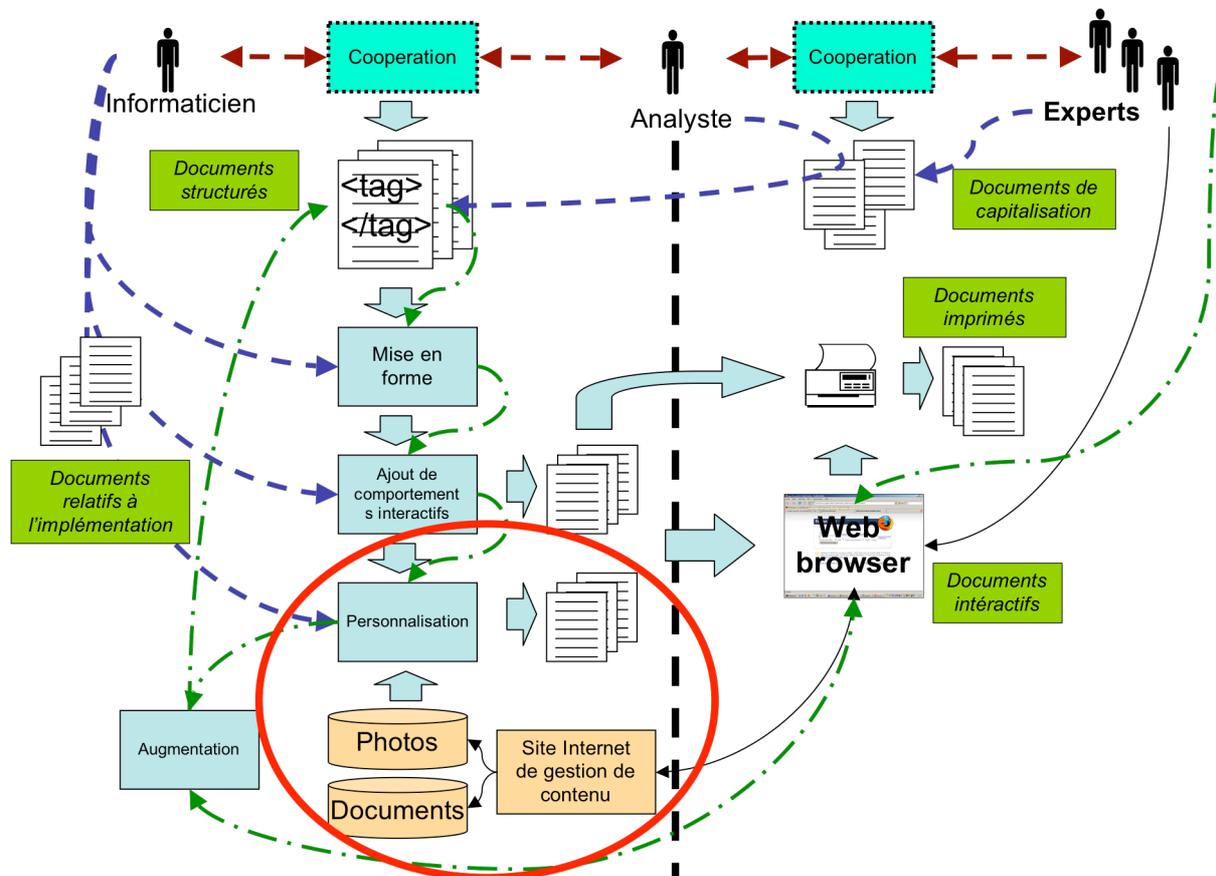


Figure 159 Intégrer la capitalisation dans le réseau social

**E - Observation du fonctionnement hydro-méca - D.1 - HM -**

**- FdV - D.1.A - HM - Fonctionnement hydro-méca de l'écluse -**

Domaine : HYDRO-MECANIQUE

**Commentaire :**

L'écoute du fonctionnement sans démontage :

- Placé à proximité des organes en mouvement (système de manoeuvre des portes, des vannes aqueduc, accouplement, pallier), il s'agit d'écouter et observer les composants en fonctionnement comme par exemple la chaîne cinématique de manoeuvre des organes etc.

**- Description -**

**Explication :**

Ecoute des phénomènes acoustiques anormaux :

- Écoute des vibrations (réducteurs, renvois d'angle, accouplements) , des bruits de grincement, des cliquetis, des frottements, des bruits de roulements etc. qui peuvent être le signe de désordres mécaniques.
- En situation normale, le fonctionnement des organes doit être fluide.
- Tout bruit anormal peut être le signe de dents casées sur les pignons, chaînes battantes, mauvais réglages ou jeu anormal, Noter les vibrations et les bruits des engrenages pendant le fonctionnement :
- Vibration des vantaux (problème de jeux ou mauvaise répartition des forces);
- Analyse vibratoire en cas de suspicion.

**Inspection :**

Portes :

- État du fonctionnement et du réglage des portes;
- Temps de manoeuvre des portes;

Mécanisme de manoeuvre des portes :

- Détection de fuite d'huile;
- Contrôler l'efficacité du graissage des crapaudines;
- Contrôler le niveau d'eau dans les puits de contrepoids;

Figure 160 Le guide avant intégration d'une zone de mise à jour

**E - Observation du fonctionnement hydro-méca - D.1 - HM -**

**FdV - D.1.A - HM - Fonctionnement hydro-méca de l'écluse -**

Domaine : HYDRO-MECANIQUE

**Commentaire :**

L'écoute du fonctionnement sans démontage :

- Placé à proximité des organes en mouvement (système de manoeuvre des portes, des vannes aqueduc, accouplement, pallier), il s'agit d'écouter et observer les composants en fonctionnement comme par exemple la chaîne cinématique de manoeuvre des organes etc.

**- Description -**

```

<content class="text">Ecoute des phénomènes acoustiques anormaux : </content>
<content class="text">- Écoute des vibrations (réducteurs, renvois d'angle, accouplements) , des bruits de grincement, des cliquetis, des frottements, des bruits de roulements etc. qui peuvent être le signe de désordres mécaniques. </content>
<content class="text">- En situation normale, le fonctionnement des organes doit être fluide. </content>
<content class="text">- Tout bruit anormal peut être le signe de dents casées sur les pignons, chaînes battantes, mauvais réglages ou jeu anormal, Noter les vibrations et les bruits des engrenages pendant le fonctionnement ;</content>
<content class="text">- Vibration des ventaux (problème de jeux ou mauvaise répartition des forces); </content>
<content class="text">- Analyse vibratoire en cas de suspicion. </content>

```

Save Cancel

**Inspection :**

Portes :

- État du fonctionnement et du réglage des portes;
- Temps de manoeuvre des portes;

Figure 161 Intégration d'une zone de mise à jour du guide lorsqu'il est consulté

## Bibliographie

- Ackerman, M. S. (1997). Communication and Collaboration from a CSCW Perspective. NSF Workshop on Human Centered Intelligent Systems, Broadcast by moderators to Red Rock Eater email distribution list (Phil Agre) and Mind, Culture, and Activity email distribution list (Michael Cole).
- Alberts, I. et S. Bertrand-Gastaldy (2006). Pratiques textuelles et genre en contexte de travail au gouvernement. Document Numérique et Société, Fribourg, Suisse, ADBS.
- Amann, B. et P. Rigaux (2003). Comprendre XSLT. Paris, O'Reilly.
- Ashby, W. R. (1956). An introduction to Cybernetics. Londres, Chapman & Hall.
- Bachimont, B. (2004). Arts et sciences du numérique : Ingénierie des connaissances et critique de la raison computationnelle Université de Technologie de Compiègne.
- Bachimont, B. (2004). "Réinterroger les structures documentaires : de la numérisation à l'informatisation." Revue I3 4(1): 54-74.
- Bachimont, B. et S. Crozat (2004). "Instrumentation numérique des documents : pour une séparation fond forme." I3 4(1).
- Baker, M. (1994). Analyse d'explications négociées pour la spécification d'un système d'aide à un système à base de connaissances.
- Baker, M. (1996). L'explication comme processus de structuration interactive des connaissances. Explication et EIAO.
- Baker, M. (2000). "Explication, Argumentation et Négociation : analyse d'un corpus de dialogues en langue naturelle écrite dans le domaine de la médecine." Psychologie de l'Interaction 9-10: 179-210.
- Baker, M. (2000). "The roles of models in Artificial Intelligence and Education research : a prospective view." International Journal of Artificial Intelligence in Education 11: 122-143.
- Baker, M., E. d. Vries, K. Lund et M. Quignard (2001). Interactions épistémiques médiatisées par ordinateur pour l'apprentissage des sciences : bilan de recherches. Actes du colloque Environnements Interactifs d'Apprentissage avec Ordinateur. 8.
- Barthès, R. (1982). L'obvis et l'obtus, essais critique III. Paris, Seuil.
- Bastien, C. (1991). Validation de critères ergonomiques pour l'évaluation d'interfaces utilisateurs, INRIA.
- Bateson, G. et Y. Winkin (1984). La nouvelle communication, Seuil.
- Bauby, C. E., B. Charbonnier, P. Haïk, S. Lacombe, J. Lonchamp et E. Remy (2005). Asset management evaluation : a pilot case study. PVP. Denver - USA.
- Bauby, C. E., P. Haïk, E. Remy et B. Ricard (2004). Knowledge modelling for asset management evaluation. PVP. San Diego.
- Beaudoin-Lafon, M. et M. Hascoët (2001). "Visualisation Interactive d'Information." REVUE I3.
- Beaudoux, O. (2004). Espaces de travail interactifs et collaboratifs : Vers un modèle centré sur les documents et les instruments d'interaction. Informatique Orsay, Université Paris XI: 226.
- Beaudoux, O. (2004). "Un modèle de composants (inter)actifs centré sur les documents." Information-Interaction-Intelligence 4(1).
- Béguin, P. et P. Rabardel (2000). "Concevoir pour les activités instrumentées." Interaction Homme-système, perspectives et recherches psycho ergonomiques, Revue IA: 35-54.
- Benard, V. (2007). Conception d'un outil pour la coopération au sein d'un réseau de santé - une approche socio-technique. Informatique. Troyes, Université de Technologie de Troyes.

- Berneers-Lee, T., J. Hendler et O. Lassila (2001) "The Semantic Web." Scientific American, DOI:
- Bertin, J. (1973). Sémiologie graphique - Les diagrammes - Les réseaux - Les cartes, Editions de l'Ecole des Hautes Etudes en Sciences.
- Bestgen, Y. et V. Dupont (2003). Impact des références intermodales sur la lecture et l'apprentissage d'un document multimédia. Actes de CIDE6: 11-24.
- Béthery, A. (2005). Guide de la classification décimale de Dewey : Tables abrégées de la XXIIe édition intégrale en langue anglaise, Broché.
- Blok, C. (1997). Dynamic visualization in a developing framework for the representation of geographic data.
- Bonin, S. (1997). Le développement de la graphique de 1967 à 1997.
- Boughzala, I. (2001). Démarche méthodologique de conception de systèmes d'information coopératifs interagents pour la gestion des connaissances. Lip 6. Paris, Paris VI.
- Boughzala, I. et J. L. Ermine (2003). Le management des connaissances. Paris.
- Boullier, D. (2001). "Les choix techniques sont des choix pédagogiques : les dimensions multiples d'une expérience de formation à distance." Sciences et Techniques Educatives.
- Bourguin, G., A. Derycke et J. C. Tarby (2005). "Systèmes Interactifs en Co évolution Réflexions sur les Apports de la Théorie de l'Activité au support des pratiques Collectives Distribuées." RIHM 6(1).
- Bouzaiène, L. (2005). Anticipation du Vieillissement par Interrogation et Simulation d'Experts. Paris, Centrale.
- Bouzaiène, L., F. Pérès, P. Haïk, E. Remy et B. Ricard (2002). Maintenance Experts Crossed Stimulation : Application to the French Nuclear Fleet. ESRL 02, Lambda Mu 13. Lyon - France.
- Briet, S. (1951). Qu'est ce que la documentation ? Paris, Edition Documentaires Industrielles et Techniques
- Buckland, M. K. (1997). "What is a "Document" ?" American Society for Information Science 48(9): 804-809.
- Casabianca, M. (2005). Extreme Programming O'Reilly.
- Castillo, O., N. Matta et J. L. Ermine (2004). De l'appropriation des connaissances vers l'acquisition des compétences Colloque C2EI, modélisation et pilotage des systèmes de connaissances et de compétences dans les entreprises industrielles, Nancy.
- Cawsey, A. (1994). "Developing an Explanation Component for a Knowledge-Based System : Discussion." Expert System with Applications 8(4): 527-531.
- Cawsey, A. (1995). "Developing an Explanation Component for a Knowledge-Based System : Discussion." Expert System with Applications 8(4): 527-531.
- Chabin, M.-A. (2004). "Document trace et document source. La technologie numérique change-t-elle la notion de document ?" Information-Interaction-Intelligence 4(1).
- Charlet, J. (2003). L'ingénierie des connaissances - Développements, Résultats et Perspectives pour la Gestion des Connaissances Médicales, Pierre et Marie Curie.
- Chateauraynaud, F. (2003). "Marlowe. Vers un générateur d'expériences de pensée sur des dossiers complexes." Bulletin de Méthodologie Sociologique(79): 5-46.
- Cotte, D. et M. Desprès-Lonnet (2004). "Le document numérique comme " lego " ou La dialectique peut-elle casser des briques ?" Information-Interaction-Intelligence 4(1).
- Cox, R. et P. Brna (1995). "Supporting the Use of External Representations in Problem Solving : the need for flexible learning environments." Journal of Artificial Intelligence in Education: 239-302.
- Crosnier, H. L. (2005). De l'édition de texte à la publication de données. Revue des rencontres internationales de Lure, Lure.

- Delalonde, C. (2007). Mise en relation et coopération dans les équipes de R&D distribuées. L'Application de l'Actor Network Theory dans la recherche de "Connaissances". Informatique. Troyes, Université de Technologie: 179.
- Dieng-Kuntz, R., O. Corby, F. Gandon, A. Giboin, J. Golebiowska, N. Matta et M. Ribière (2001). Méthodes et outils pour la gestion des connaissances : Une approche pluridisciplinaire du Knowledge Management (2ème édition), Dunod.
- Dourgnon-Hanoune, A., P. Salaün et C. Roche (2006). Ontology for Long-Term Knowledge. IEA / AIE, Springer Berlin / Heidelberg.
- Dourish, P. (1999). Where the action is - The Foundations of Embodied Interaction, MIT Press.
- Dufresne, A. (1997). "Conception d'interfaces pour l'apprentissage à distance." La revue de l'éducation à distance **12**: 177-200.
- Dufresne, A. (2001). "Conception d'une interface adaptée aux activités de l'éducation à distance - ExploraGraph." Sciences et Techniques Educatives.
- Eco, U. (1988). Le Signe.
- Eisenberg, J. D. (2003). SVG : Production orientée XML de graphiques vectoriels. Paris, O'Reilly.
- El-Hachini, M. (2005). Indexation des documents multilingues d'actualités incluant l'arabe : équivalence interlangues et gestion des connaissances chez les indexeurs. Sciences de l'information et de la communication. Lyon, Université Lumière (Lyon 2).
- Enjalbert, P. (1996). "De l'interprétation (sens, structure et processus)." Intelecta.
- Ermine, J. L. (1996). Les systèmes de connaissances, Hermès.
- Ermine, J. L. (2000). La gestion des connaissances un levier stratégique pour les entreprises. Actes de IC 2000. Toulouse.
- Ermine, J. L. (2003). La gestion des connaissances, Hermès science publications.
- Ermine, J. L., M. Chaillot, P. Bigeon, B. Charreton et D. Malavieille (1997). MKSM, a method for knowledge management. 5th International Symposium on the Management of Industrial and Corporate Knowledge, Compiègne, France.
- Ertzscheid, O. (2002). Les enjeux cognitifs et stylistiques de l'organisation hypertextuelle. Toulouse, Université de Toulouse II - Faculté des lettres et sciences humaines.
- Ertzscheid, O. (2006). Indexation sociale et continents documentaires. Document numérique et société (DocSoc), Fribourg - Suisse.
- Falquet, G. (1999). Des documents virtuels pour lire les bases de données Des documents virtuels pour lire les bases de données Des documents virtuels pour lire les bases de données. Montpellier.
- Gaglio, G., M. Maroccia et M. Zacklad (2006). L'usage des diapositives numériques en milieu organisé ou la recombinaison provisoire de collectifs éclatés. Pratiques et usages organisationnels des sciences et technologies de l'information et de la communication, Rennes.
- Gagnière, J. (2008). Guide des vannes (en cours de rédaction). Chatou, EDF R&D.
- Gagnon-Arguin, L. (1992). L'archivistique: Son Histoire, Ses Acteurs Depuis 1960, Archives Study and teaching Quebec: 250.
- Gehtland, J., B. Galbraith et D. Almaer (2006). Ajax par la pratique. Paris, O'Reilly.
- Gibson, J. (1977). "The Theory of Affordances." Perceiving, Acting, and Knowing.
- Gibson, J. (1979). "The Ecological Approach to Visual Perception."
- Goody, J. (1979). La raison Graphique - la domestication de la pensée sauvage, Les Editions de Minuit
- Grassaud, A. (2000). La formation des personnels dans un milieu industriel à "haut risque technologique.

- Grassaud, A. et S. Parfouru (2007). Guide générique d'aide au diagnostic des écluses version V1 & REX (H-P1B-2007-01494-FR). Chatou, EDF R&D.
- Grenier, C. et B. Pauget (2006). "Analyse de la création de connaissances métier dans des réseaux d'acteurs professionnels : le rôle de la connaissance relationnelle." Gestion 2000 4(06).
- Grunstein, M. et J. Barthès (1996). AN Industrial View of the Process of Capitalizing Knowledge.
- Grunstein, M. et J. P. Barthès (1996). An Industrial View of the process of Capitalizing Knowledge. Fourth International ISMICK Symposium Proceedings, Rotterdam, Netherlands, ERGON VERLAG.
- Gurwitsch, A. (1935). Développement historique de la Gestalt-Psychologie, Thalès.
- Haïk, P. (2004). Affaire A002W - "Veille Technologique et Partenariats en Traitement de l'Information" : Synthèse sur la Stimulation d'Experts pour l'anticipation de Défaillances (TRIZ, AFD, etc.). Paris EDF R&D.
- Haïk, P. (2005). Projet " CapCoV " - Spécification Fonctionnelle de l'Algorithme de Recherche des Configurations (Matériaux, Milieux, Sollicitations) Pertinentes pour un Composant. (H-P1B-2005-02094-FR). Chatou, EDF R&D.
- Haïk, P., S. Mahé et B. Ricard (2002). Knowledge engineering as a support for decision making in plant operation and maintenance. Tokyo, Japan.
- Haïk, P., S. Mahé et B. Ricard (2003). Knowledge Modeling as a Support for Knowledge Acquisition, Sharing, Operationalization and Maintenance. Actes de Icone 11.
- Haïk, P. et S. Parfouru (2008). "Adaptative" User Interface as a Support for Evaluation Analysis in the Context of Asset Management. ASME PVP, Chicago, USA.
- Haïk, P., S. Parfouru, C. Bauby et S. Mahé (2007). "Adaptative" User Interface for Expert Knowledge Validation and Evaluations Interpretation in an Asset Management Process. ASME PVP, San Antonio, Texas.
- Hatchuel, A. et B. Weil (1992). L'expert et le système, suivi de quatre histoires de systèmes-experts. Paris.
- Hearst, M. (1995). TileBars: Visualization of term distribution information in full-text information access. Human Factors in Computing Systems, ACM.
- Henry, N. et J.-D. Fekete (2006). MatrixExplorer: Un système pour l'analyse exploratoire de réseaux sociaux. IHM, Montréal, ACM.
- Hutchins, E. (1991). The social organization of distributed cognition. Perspectives on Socially Shared Cognition. J. M. L. S. D. T. L. B. Resnick
- J. Ellson, Emden Gansner, L. Koutsofios, Stephen North et G. Woodhull (2003). Graphviz and Dynagraph - Static and Dynamic Graph Drawing Tools Graph Drawing Software. M. J. a. P. Mutzel, Springer-Verlag: 127-148.
- Jean-Paul Barthès, Rose Dieng et G. Kassel (1999). "Mémoire d'Entreprise." bulletin de l'AFIA(n°36).
- Jeanneret, Y. et C. Tardy (2007). Métamorphoses médiatiques, pratiques d'écriture et médiation des savoirs. Paris, LaLICC - Paris 4.
- Kaptelinin, V. (2003). "Learning with artefacts : integrating technologies into activities." Interacting with Computer 15: 831-836.
- Karouach, S. et B. Dousset (2003). "Visualisation de relations par des graphes interactifs de grande taille." Journal of ISDM (Information Sciences for Decision Making) 6(57): 12.
- Knuth, D. (1968). The Art of Computer Programming, Addison-Wesley.
- Knuth, D. (1984). "Literate Programming." The Computer Journal 27(2).
- Kovacs, B., F. Gaunet et X. Briffault (2004). Les techniques d'analyse de l'activité pour l'IHM, Hermès Science.

- Krasmer, G. E. et S. T. Pope (1988). "A cookbook for using the Model-View-Controller user interface paradigm in Smalltalk-80." Journal of Object Oriented Programming 26)27.
- L'Hédi Zaher, Jean-Pierre Cahier et M. Zacklad (2007). De la recherche d'information à la recherche ouverte d'information. IEEE Conference on Sciences of electronics, Technologies of Information and Telecommunication (SETIT'2007), Tunisie.
- Lacomme, P., C. Prins et M. Sevaux (2003). Algorithmes de graphes.
- Lamy, J. B. (2006). Conception et évaluation de méthodes de visualisation des connaissances sur le médicament - Mise au point d'un langage graphique et application aux connaissances sur le médicament. Informatique médicale Paris, Paris 6
- Le Blanc, B., C. Obernesser et B. Claverie (2003). "Validation et évaluation cognitives de techniques de navigation et de visualisation de données." Revue des sciences et technologies de l'information, RSTI série RIA-ECA (1-2-3)(17): 93-104.
- Le\_Coadic, Y.-F. (1994). La Science de l'information, PUF.
- Lengler, R. et M. Eppler (2007). Towards A Periodic Table of Visualization Methods for Management. Graphics and Visualization in Engineering (GVE 2007), Clearwater, Florida, USA., IASTED.
- Leplat, J. (2004) "Elements pour l'étude des documents prescripteurs." activités 1, DOI:
- Lortal, G., M. Lewkowicz et A. Todirascu-Courtier (2007). Des activités d'annotation : De la glose au document. Annotation dans les documents pour l'action. P. Salambier, Hermès
- Lux-Pogodalla, V. et J.-Y. Vion-Dury (2004). "Réflexions sur la modélisation des documents." Information-Interaction-Intelligence 4(1).
- Mahé, S. (2004). Projet CapCoV - Spécification Technique et Fonctionnelle Détaillée d'une Maquette d'Outil de Capitalisation des Connaissances sur les Mécanismes de Vieillesse. (H-P14-2004-03078-FR) Chatou, EDF R&D.
- Mahé, S. et S. Parfouru (2006). Projet "XXX" - Restitutions et parcours des bases de connaissances XXX et du corpus documentaire (H-P1B-2006-04490-FR). Chatou, EDF R&D.
- Mallein, P. et S. Tarozzi (2002). "Des signaux d'usage pertinents pour la conception des objets communicants." Les cahiers du numérique 3(4).
- Malvache, P. et P. Prieur (1993). Mastering corporate experience with Rex method. ISMICK, Compiègne, France.
- Marcoux, Y. (2006). A natural-language approach to modeling Why is some XML so difficult to write. Extreme Markup Language, Montréal.
- Marcoux, Y. et C. Rémillard (2001). Cooperative Work One Step Forward: How XML Registries Will Help Dissolve the Boundary between Developer and User Communities.
- Mas, S. (2004). Impact de l'organisation des documents électroniques sur l'interprétation de l'information organique et consignée dans un contexte de gestion décentralisée. Le numérique : impact sur le cycle de vie du document. Université de Montréal.
- Matta, N. (2004). Ingénierie des connaissances en conception pour la mémoire de projets. H. à. d. d. Recherches, Université de technologie de Compiègne.
- Morand, B. (2000). Le diagramme à la périphérie ou au coeur de la cognition. Rencontres interdisciplinaires sur les représentations graphiques dans les systèmes complexes naturels et artificiels, 9èmes Journées de Rochebrune, .
- Morand, B. (2004). Logique de la conception - Figures de sémiotique générale d'après Charles S. Pierce, L'Harmattan.
- Mugnier, M.-L. et M. Chein (1995). "Représenter des Connaissances et Reasonner avec des Graphes." RIA 10.

- Nanard, J. (1990). La manipulation directe en interface homme-machine. Montpellier, Montpellier II.
- Nanard, M. et J. Nanard (2005). Saphir : un cadre de référence pour une spécification des hypermédias par intention. Hypertexte et hypermédia (H2PTM) - Créer, Jouer, Echanger. Paris.
- Newell, A. (1981). "The Knowledge level." AI Magazine 2(2): 1-20.
- Nicolle, A. (2003). Interaction langagière personnes / machines. Variation, construction et instrumentation du sens, Hermès: 251-285.
- Nicolle, A. et V. S. Dizier (1999). Vers un modèle des interactions langagières. L'interdisciplinaire. Lyon.
- Nigay, L. et J. Coutaz (1993). ESPACE PROBLEME, FUSION ET PARALLELISME DANS LES INTERFACES MULTIMODALES. L'interface des Mondes Réels et Virtuels. Montpellier.
- Nonaka, I. et H. Takeuchi (1995). The Knowledge Creating Company, Oxford University Press.
- Norman, D. A. (1991). "Cognitive artefacts." Designing interaction.
- Norman, D. A. (1999). "Affordances, Conventions and Design." Interactions 1(3): 38-42.
- O'Neill, E. (2001). User-Developer Cooperation in Software Development - Building Common Ground and Usable Systems. London.
- Oliveira Prates, R., C. S. De Souza et A. Cristina Bicharra Garcia (1997). "A semiotic framework for multi-user interfaces." SIGCHI 29(2).
- Parfouru, S., A. Grassaud, S. Mahé et M. Zacklad (2006). Document pour l'Action comme media pour la Gestion des Connaissances CIDE 9. Europia: 199-215.
- Parfouru, S. et P. Haïk (2007). Approche de Capitalisation et Restitution sur un Système à base de connaissances existant (H-P1B-2007-03398-FR) - en cours de finalisation. Chatou, EDF R&D.
- Parfouru, S., P. Haïk, S. Mahé et M. Zacklad (2007). Analyse de la construction de documents de capitalisation de connaissances au travers de documents structurés : vers le développement d'un atelier sémiotique de composition de documents. CIDE 10, Nancy.
- Parfouru, S., S. Mahé et M. Zacklad (2006). Dispositif Documentaire pour la Gestion des Connaissances. Conception et Innovation (CONFERE). Marakech (Maroc).
- Parfouru, S., S. Mahé et M. Zacklad (2007). Using Electronic Document as an Interface for Knowledge Based System. Workshop on Knowledge Management and Organizational Memories - International Joint Conference on Artificial Intelligence. Hyderabad - India.
- Pédauque, R. T. (2006). Le document à la lumière du numérique. Caen, C&F Editions.
- Quint, V. et I. Vatton (2004). Techniques for AAuthoring Complex XML Documents. Document Engineering, Milwaukee ACM Press.
- Rastier, F. (1987). Sémantique interprétative, PUF.
- Ray, E. T. (2001). Introduction à XML Paris, O'Reilly.
- Ricard, B. (2007). Proposition d'évolution des fonctions de restitution des résultats de diagnostic de l'outil µBULLES (H-P1B-2007-01501-FR), EDF R&D.
- Ricard, B., P. Haïk et J. Lambert (2000). Un atelier pour la capitalisation et l'exploitation de connaissances de maintenance d'équipements industriels: 36-44.
- Rose Dieng, Olivier Corby, Alain Giboin et M. Ribière (98). Methods and Tools for Corporate Knowledge Management. 11th Workshop on Knowledge Acquisition, Modeling and Management, Banff, Canada.

- Rousseaux, F. et T. Bouaziz (2006). "Comment, sans nuire aux informaticiens, dépasser l'artifice des Connaissances ? Documents numériques, Connaissances et Collections." Document Numérique 9(1): 93-108.
- S. Pook, E. Lecolinet, G. Vaysseix et E. Barillot (2000). Control Menus: Execution and Control in a Single Interactor. CHI, The Hague, The Netherlands, ACM Press.
- Saad, I., C. Rosenthal-Sabroux et M. Grundstein (2005). Une étude approfondie pour le choix des connaissances à capitaliser en amont de la construction d'une mémoire d'entreprise. Ingénierie des connaissances, Nice.
- Saleh, I. (2005). Les hypermedias - conception et réalisation
- Schmidt, K. et C. Simone (1996). "Coordination mechanisms: Towards a conceptual foundation of CSCW systems design. ." CSCW Journal, 5(2-3): 155-200.
- Searle, J. R. et D. Vanderveken (1985). Foundations of Illocutionary Logic. Cambridge, Cambridge University Press.
- Searle, J. (1972). Les actes de langage, Essai de philosophie du langage, Hermann.
- Shneiderman, B. (1992). "Shneiderman, B. (1992). Tree Visualization with Tree-Maps: 2-D space-Filling Approach. ." ACM Transactions on Graphics 11(1): 92-99.
- Shneiderman, B. (1996). "The Eyes Have It : A Task by Data Type Taxonomy for information Visualizations." 336-343.
- Simon, H. A. (1978). "Rationality as Process and as Product of Thought." American Economic Review 68(2): 1-16.
- Smith, J. B. (1994). Collective Intelligence in Computer-Based Collaboration, Lawrence Erlbaum.
- Soulier, E. (2003). Technique de Storytelling pour le partage de connaissances dans les communautés de pratique. Informatique. Paris, Université Pierre et Marie Curie (Paris VI): 397.
- Souza, C. S. d. (1996). The Semiotic Engineering of Concreteness and Abstractness: From User Interface Languages To End User Programming Languages.
- Stuart K. Card, Jock D. Mackinlay et B. Shneiderman (1999). Readings in Information Visualization: Using Vision to Think. San Francisco, Morgan Kaufmann
- Tazi, S., K. Drira et K. Essajidi (2006). Maintien de la cohérence des intentions de communication dans la rédaction coopérative. Conférence Internationale sur la Document Electronique (CIDE) Fribourg (Suisse), Europa.
- Trognon, A. et C. Brassac. (1992). " L'enchaînement conversationnel." Cahiers de Linguistique Française 13: 76-107.
- Tufte, E. R. (1997). Visual and Statistical Thinking: Displays of Evidence for Making Decisions, Graphics Press.
- Wenger, E. (2001). Communities of Practice : Learning as a social system, Cambridge University Press.
- Wikipedia Wikipédia, projet d'encyclopédie librement réutilisable que chacun peut améliorer.
- Zacklad, M. (2000). Ingénierie des connaissances appliquée aux systèmes d'information pour la coopération et la gestion des connaissances, Paris VI.
- Zacklad, M. (2004). Processus de documentation dans les Documents pour l'Action (DopA) : statut des annotations et technologies de la coopération associées. Le numérique : Impact sur le cycle de vie du document pour une analyse interdisciplinaire, Montréal.
- Zacklad, M. (2005). Introduction aux ontologies sémiotiques dans le Web Socio Sémantique. Nice.
- Zacklad, M. (2007). Annotation : : attention, association, contribution. Annotation dans les documents pour l'action. P. Salambier, Hermès.

- Zacklad, M. (2007). L'économie de fonctionnalité encastrée dans la socio-économie des transactions coopératives : dynamique servicielle et fidélisation soutenable. L'économie des services pour un développement durable, Cerisy.
- Zacklad, M., J.-P. Cahier et X. Pétard (2003). Du Web cognitivement sémantique au Web socio sémantique - Exigences représentationnelles de la coopération. Journée "Web sémantique et Sciences humaines et sociales » organisée par l'action spécifique "Web sémantique » CNRS/STIC.
- Zacklad, M., N. Legay et J. Blanquet (1996). La modélisation conceptuelle des IHM avec PSS. actes du 5ème Colloque Ergo-IA: 101-115.

