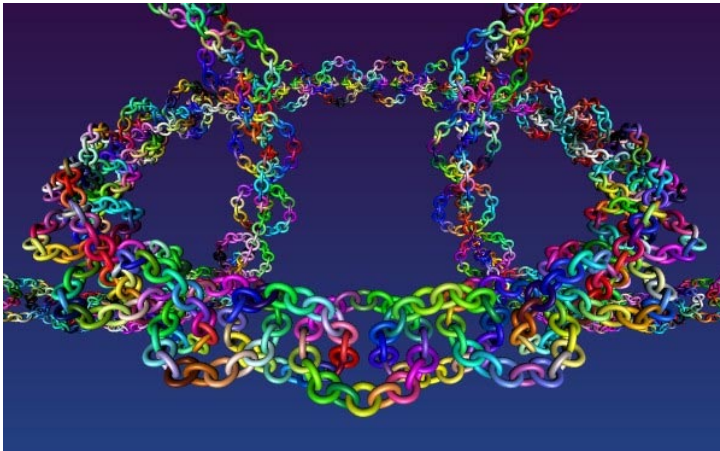


# CONCLUSION



Organisation hypertextuelle, pragmatique de la connaissance & écologie cognitive.

*« Un événement s'est produit en effet dans tous les domaines de la connaissance qui a dérivé vers des horizons nouveaux la science contemporaine : la connaissance a cessé de se déployer dans le seul élément du savoir pour devenir recherche ; en d'autres termes, elle s'est détachée de la sphère de la pensée où elle trouvait sa patrie idéale pour prendre conscience d'elle-même comme cheminement à l'intérieur d'un monde réel et historique où se totalisent techniques, méthodes, opérations et machines. La science n'est plus un chemin d'accès à l'énigme du monde, mais le devenir d'un monde qui ne fait plus maintenant qu'une seule et même chose avec la technique réalisée. En cessant d'être seulement savoir pour devenir recherche, la science disparaît comme mémoire pour devenir histoire ; elle n'est plus une pensée mais une pratique, non plus un cycle fermé de connaissances, mais, pour la connaissance, un chemin qui s'ouvre là même où il s'arrête. »* [Foucault 94 p.155]

*« C'est plus qu'un bon mot quand Lévi-Strauss baptise la science de l'homme entropologie. »* [Steiner 98 p.225]

## 7.4.2. Sérendipité structurelle.

Admettons pour l'exemple, que nous nous trouvions dans une bibliothèque, à la recherche d'une thèse déjà repérée, pour construire un état de l'art sur une question donnée. Non loin de la thèse recherchée, sur le même rayonnage, figure une autre thèse dont le titre est évocateur et dans laquelle, après lecture rapide, nous trouvons effectivement des informations intéressantes. La sérendipité ici à l'œuvre est de type structurelle : elle est liée à une identification, à un parallélisme formel, structurel (de fait on est dans le rayonnage des thèses et non dans celui des journaux qui eût été moins approprié pour l'objectif de notre recherche initiale).

Admettons maintenant que nous effectuions la même recherche, dans la même bibliothèque, mais cette fois en consultant l'une des bases de données dont elle dispose : on utilisera alors les champs structurés de la base de donnée pour exprimer notre requête (mots du titre, nom de l'auteur, etc.). Selon la règle de « matching » applicable à tout type d'information structurée, l'échelle du phénomène de sérendipité se réduit considérablement, même si elle reste possible (un même auteur ayant pu rédiger deux thèses différentes par exemple) et demeure de nature structurelle.

## 7.4.3. Sérendipité associative.

Sur Internet, et plus généralement dans tout système distribué d'information non-structurée, ce phénomène change de nature et se donne à lire avec une acuité déterminante dans les stratégies de navigation choisies par les utilisateurs. Si l'on interroge un moteur de recherche en entrant une série de mots-clés (qui peuvent être les mêmes que ceux utilisés pour l'interrogation de la base de donnée), deux cas se présentent :

- le moteur de recherche dispose, dans sa base de donnée ou dans sa base d'index, d'informations présentant un relatif niveau de structuration (c'est par exemple le cas des annuaires de recherche si on les interroge en utilisant les catégories qu'ils proposent) : le phénomène de sérendipité structurelle reste opérant. Au vu du nombre de résultats possibles, dans ce cas comme dans les deux premiers évoqués (interrogation du rayonnage des thèses ou d'une base de donnée), le facteur déterminant consiste à limiter le silence (absence de résultats) ;
- le moteur de recherche ne dispose pas d'information structurée – ce qui demeure le cas le plus fréquent – et les listes de résultats qu'il présente à la requête de l'utilisateur sont alors considérables. La sérendipité se manifestant cette fois dans l'affichage possible d'un résultat pertinent bien que ne correspondant pas aux termes exacts de la requête est alors de nature associative. Le facteur déterminant dans les stratégies de navigation qui seront alors mises en place par l'internaute est celui qui lui permettra d'éviter le bruit et non plus le silence.

Notons ici que la sérénité associative résulte de la conjugaison de phénomènes sémantiques, algorithmiques, individuels (usages) et techniques (référencement, balises méta<sup>235</sup>, spam<sup>236</sup> ...). On peut par ailleurs constater, avec la dernière génération de moteurs de recherche que le facteur déterminant redevient celui du modèle classique, c'est-à-dire éviter le silence<sup>237</sup>.

Cette sérénité a tout au moins comme mérite méthodologique d'établir qu'il n'est pas nécessairement plus facile de trouver de l'information dans un système ordonné, structuré et formaté que, comme cela semble être le cas pour le web, dans un système d'information caractérisé par une forte entropie et ne disposant en tout cas d'aucun niveau de contrôle unique.

« (...) l'aspect chaotique et dynamique du web est une bénédiction. De tous les médias qu'il y a eu avant lui, le web est celui qui reflète le mieux notre forme de pensée. Tandis qu'il croît et se complexifie, rappelons nous que l'homonymie et l'ambiguïté font partie de nos formes d'expression et de pensée. Et nous avons nous aussi besoin de liens incongrus pour être créatifs : c'est peut-être même la caractéristique la plus importante chez l'humain. Alors apprenons à apprivoiser le fortuit. » [Feat 97 p.163]

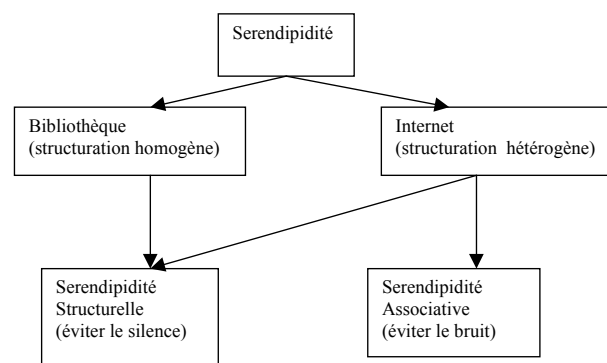


Fig. 13 : Sérénité structurelle et associative.

## 7.5. Comment lutter contre ces phénomènes ?

Dans un article fondateur déjà évoqué à de nombreuses reprises dans ce travail, [Balasubramanian 94] dresse un panorama exhaustif des solutions envisagées pour réduire la désorientation et surcharge cognitive. Nous nous proposons ici de reprendre cette liste en la commentant à la lumière des conclusions que notre étude a jusqu'ici permis d'établir.

<sup>235</sup> en HTML, ces balises permettent aux auteurs de contrôler l'indexation de leurs documents.

<sup>236</sup> le « spam » désigne les pratiques frauduleuses qui permettent de « fausser » l'indexation d'un document (faux mots-clés ...)

<sup>237</sup> certaines pratiques sont à ce titre tout à fait éclairantes du point de vue d'une « sociologie » de la recherche d'information, comme celle du « GoogleWhacking » (<http://www.googlewhack.com>) : Google étant le moteur de recherche le plus en vogue et celui disposant de la plus grande base d'index, cette pratique consiste à formuler des requêtes ne ramenant qu'une seule réponse.

## Citations originales.

### - Point 6. FoRSIC.

- [Illich 76 p.80] « The planning of new educational institutions (...) must not start with the question, 'What should someone learn ?' but with the question, 'What kind of things and people might learners want to be in contact with in order to learn ?' »
- [Brown et al. 89] « People who are not taking part directly in a particular activity learn a great deal from their legitimate position on the periphery. »
- [Brown & Duguid 92] « it seems important not simply to fragment or decompose tasks to make them didactically tractable on their own and for individuals. Any decomposition of the task must be done with an eye not to the task or the user in isolation, but to the learner's need to situate the decomposed task in the context of the overall social practice. The presence of the full context gives the learner the chance to « steal » whatever he or she finds most appropriate. It is vitally important not to fragment the social periphery. »

### - Point 7. Le rôle à jouer des ontologies.

- [Sowa 01] « The art of ranking things in genera and species is of no small importance and very much assists our judgment as well as our memory. (...) This helps one not merely to retain things, but also to find them. »
- [Gruber 93] « An ontology is a formal, explicit specification of a shared conceptualization. »
- [Mizoguchi et al. 96] « system of primitive vocabulary / concepts used for building artificial systems. »
- [Carr et al. 01] « marking up is no good unless everyone speaks the same language. »
- [Carr et al. 01] « using a common vocabulary is no good unless everyone means the same thing. »
- [Mizoguchi et al. 96] « a theory of labels of concepts. »
- [Mizoguchi et al. 96] « vocabulary is language-dependent, so it lacks the universality. Vocabulary is weak in description of the relations among the terms in it. »
- [Mizoguchi et al. 96] « Taxonomy is often a hierarchy of concepts in which each link is a is-a link or a part-of link. »
- [Carr et al. 01] « Most ontologies have three major components that can be used in inference : a taxonomie, relationships between concepts and axioms (rules). »
- [Carr et al. 01] « provide a shared and common understanding of a domain that can be communicated across people and applications. »
- [Carr et al. 01] « they all include a vocabulary of terms and some specification of the meaning of the terms. »
- [Sowa 01] « the difference between a terminological ontology and a formal ontology is one of degree : as more axioms are added to the terminological ontology, it may evolve into a formal (...) ontology. »
- [Mizoguchi et al. 96] « Task ontology is a system / theory of vocabulary for describing inherent problem solving structure of all existing tasks domain – independently. It is obtained by analysing task structures of real world problems. Design of task ontology is done in order to overcome the shortcomings of generic tasks while preserving their basic philosophies. The ultimate goal of task ontology research includes to provide theory of all the vocabulary necessary for building a model of human problem solving process.
- Task ontology consists of the following four kinds of concepts :
  - generic nouns representing objects reflecting their roles appearing in the problem solving process
  - generic verbs representing unit activities appearing in the problem solving process
  - generic adjectives modifying the objects
  - other concepts specific to the task. »
- [Illich 76 p.81]
  - « Reference Services(...) – which facilitate access to things or processes used for formal learning. (...)
  - Skill exchanges – which permit persons to list their skills, the conditions under which they are willing to serve as models for others who want to learn these skills, and the addresses at which they can be reached.
  - Peer-matching – a communications network which permits persons to describe the learning activity in which they wish to engage, in the hope of finding a partner (...).
  - Reference Services to Educators-at-large – who can be listed in a directory giving the addresses and self-descriptions of professionals, paraprofessionals and freelancers, along with conditions of access to their services. »
- [Carr et al. 01] « Some conceptual hypermedia systems expose the ontology and make it explicitly navigable ; others make the classification scheme more implicit. Should the ontology be visible during linking ? When the ontology becomes a sophisticated model of roles and axioms rather than a simple static tree, choosing a concept becomes an expression construction exercise. »

différents types ontologiques utilisés dans le dispositif et permet, non seulement de les croiser pour produire des ressources de formation, mais également de fonder l'ensemble du parcours permettant de créer une nouvelle ressource sur le réseau de liens qui relie les différents formalismes. Il est ainsi possible de naviguer à partir de tout ou partie d'une ressource existante vers d'autres qui sont reliées :

- par la sémantique des concepts qu'elles exploitent (modèle de domaine),
- par les niveaux de formation auxquels elles se réfèrent,
- par l'intention pédagogique de l'auteur (elle même liée aux activités cognitives nécessaires),
- par la catégorie de cours que l'on souhaite créer (correspondant à la classe « délivrable » dans la matrice de compétences),
- par l'objectif pédagogique visé par la ressource (celui-ci pouvant être associé aux différentes catégories de cours),

L'ensemble de la navigation parmi ces formalismes, s'effectue de manière transparente pour l'utilisateur qui peut les utiliser de manière conjointe ou individuellement. La ressource produite en sortie du dispositif est codée en XML et reprend l'ensemble de ces paramètres dans son balisage, autorisant du même coup autant de recherches possibles (par intention, par concepts, par catégorie, ...). SABRE intègre ainsi au maximum de leur potentiel des métadonnées souvent soulignées comme essentielles dans le cadre de l'ingénierie documentaire ou du web sémantique, mais souvent développées selon des procédures sans lien avec celles se rapportant aux ontologies (figure 20).

ForSIC n'entre pas vraiment dans la catégorie des outils de type groupware même s'il est centré sur des procédures de coopération. Il tiendrait plutôt du « communityware »<sup>99</sup>. Nous préférons le définir comme un espace collaboratif d'apprentissage (CLS)<sup>100</sup>. La collaboration est sa modalité, l'apprentissage (la formation) sa finalité, et l'espace sa condition d'existence (cartographie, arbres connaissance, organisation hypertextuelle) : dans un CLS, les modes de production et d'organisation du savoir et de la connaissance d'une part (tacite, implicite, explicite, individuelle, inter-personnelle, collective ...), et les liens qui se tissent et que tissent les représentations sociales d'autre part, sont interdépendants, co-spécifiés et co-déterminés selon une boucle de feedback dynamique sans cesse reproduite (correspondant à la spirale du processus de conversion de connaissance).

La place ainsi accordée à un ensemble de modalités de liaison (incluant le lien social), la combinaison sous un mode d'organisation hypertextuel de modèles formels de haut niveau et de niveaux de collaboration hybrides (conjuguant humain et machine) et clairement identifiés, nous permet de parler à son sujet d'un modèle pérenne pour une pragmatique de la connaissance.

<sup>99</sup> néologisme apparu en 1998 à Kyoto lors d'une conférence sur ce thème.

<sup>100</sup> CLS : Collaborative Learning Space.

Pour ne pas alourdir notre propos et parce qu'elles reposent sur des principes établis et couramment utilisés, nous n'aborderons pas ici les pratiques recommandées par la nettiquette<sup>238</sup>, ni celles relevant exclusivement de procédures techniques de gestion de sites<sup>239</sup>.

### **7.5.1. Optimiser les procédés de liaison.**

Nous considérons que l'utilisation des invariants dégagés pour les ancrs hypertextuelles<sup>240</sup> constitue de fait la première solution à ces problèmes et n'y revenons pas dans cette partie. Nous indiquons qu'elles figurent pour partie dans l'inventaire de [Balasubramanian 94] recommandant l'utilisation de :

- menus emboîtés (« *embedded menus* »)<sup>241</sup>
- notes de bas de page, historique de navigation, échelles de temps (« *Backtracking, Timestamps* »)
- « signets », nœuds les plus visités et graphes complets.

De fait, l'ensemble de ces techniques courantes censées réduire la désorientation (quand elles ne l'accroissent pas par un usage inapproprié) peut être scindé en deux ensembles distincts : le premier est celui qui concerne les fonctionnalités de typage des ancrs et vise à « *offrir une vue locale des rôles individuels de chaque lien mais ne donne que peu d'indications sur la manière dont l'ensemble auquel appartiennent les liens est structuré.* » [Young 90 p.240] Le second est celui qui concerne plus spécifiquement la structuration des nœuds : « *[les nœuds structurés] aident les utilisateurs à reconnaître et à interpréter ce qui est contenu dans le nœud, et à prendre des décisions sur quoi faire après l'avoir parcouru et comment le faire plus facilement. (...) Mais la structure interne des nœuds ne donne souvent que peu d'indications sur la manière dont une collection de nœuds est reliée.* » [Young 90 p.239] Dans l'un et dans l'autre cas et même en se plaçant dans un contexte idéal où toute sur-utilisation de l'un ou l'autre de ces procédés serait exclue, il apparaît évident que les gains fournis ne sauraient suffire à compenser les pertes occasionnées. Il ne s'agit que d'une étape qui doit être complétée par un travail sur la structure et l'organisation interne de l'information affichée en termes de cohérence sémantique et contextuelle.

<sup>238</sup> il s'agit de s'adapter à l'utilisateur en tenant compte d'aspects techniques (temps d'accès aux pages, problèmes de compatibilité ...) ou plus liés aux contenus (phénomène du temps de latence, accès hiérarchisé et différencié à l'information, chartes graphiques ...)

<sup>239</sup> il s'agit ici pour l'essentiel du « tracking » et du « profiling ». [LAB[au] 01] : « *Le tracking opère comme un dispositif qui collecte des informations sur des utilisateurs de manière « statique », de la configuration matérielle de leurs ordinateurs jusqu'aux informations personnelles, ou d'une manière « dynamique » en suivant à la trace leurs évolutions dans le réseau. (...) C'est le mapping des utilisateurs à l'échelle du groupe. (...) Le profiling consiste (...) à extraire le sens produit par les interactions de l'utilisateur dans le but de comprendre ses comportements. Il s'agit du mapping des utilisateurs à l'échelle individuelle, de la construction d'un profil sociologique jusqu'à un profil psychologique, et du reformatage de l'information en fonction de ce profil. En effet, le profiling inclut des éléments pro-actifs comme l'affichage « personnalisé » de certaines informations prioritaires par rapport à son intérêt plutôt que d'autres informations généralistes.* »

<sup>240</sup> propriétés individuelles, liées au contexte et collectives.

<sup>241</sup> mal utilisés, ces menus emboîtés peuvent avoir l'effet inverse puisque l'on se retrouve potentiellement avec trois niveaux distincts et différents de navigation possibles : celui du navigateur (boutons back et forward), celui du site en lui-même, et celui des menus de type « pop-up ».

### **7.5.2. Organiser l'architecture de contenu.**

Là encore, nous avons déjà proposé des solutions qui figurent essentiellement dans les propriétés collectives des ancrs et l'utilisation de différents types de structures (arborescentes ou non-arborescentes). Nous voulons ici remettre en perspective ces propositions pour dégager deux niveaux d'application distincts de ces structures à l'organisation sémantique des contenus, ces deux niveaux ayant des implications distinctes sur la navigation.

Le premier est celui que nous qualifierons de sémantique « interne » qui ne s'intéresse qu'au contenu sémantique propre des pages dans la linéarité temporelle d'une session de navigation (niveau sémantico-temporel), le second est celui de la sémantique « externe » permettant de gérer l'échelle des rapports possibles entre l'espace « fenêtré » d'une navigation et l'influence en termes de feedback de ce fenêtrage sur les contenus affichés. Ce second niveau peut alors être qualifié de « spatio-sémantique ». Précisons ici que bien que relevant en première main d'une intentionalité de type auctoriale, le fait de privilégier des repères temporels ou spatiaux de navigation aura des effets différents voire opposés sur des utilisateurs dont les styles cognitifs usuels privilégient l'un ou l'autre de ces repères.

#### **7.5.2.1. Sémantique interne (niveau sémantico-temporel).**

Le premier moyen d'action concernant la sémantique interne des contenus est celui hérité du codex qui, au travers de l'affichage d'un sommaire ou d'une table des matières, permet une lecture tabulaire et non plus linéaire. Dans le cadre de l'hypertexte, c'est alors l'utilisation de liens de renvoi ou liens « tabulaires » qui sera privilégiée, la navigation s'effectuant par la consultation des sections et sous-sections fondamentales d'un site au travers de cette table des matières. Si pour ce point précis, les possibilités offertes par l'hypertexte ne révolutionnent en rien les principes de lecture hérités de l'âge de l'imprimerie, et si bien au contraire elles ont tout intérêt à les reprendre à leur compte, ces *habitus* doivent cependant être pensés en termes de « page-écran », en intégrant les contraintes<sup>242</sup> liées à ce nouveau support. L'un des principes qui exemplifie le mieux cette transposition de codes liés au codex vers des implémentations intégrant les contraintes d'une page-écran est celui que nous appellerons le principe de double orientation, qui peut être schématisé comme suit :

<sup>242</sup> contraintes : gabarit, espace de sécurité (celui restant une fois éliminés les ascenseurs, les boutons du navigateur ...), affichage dynamique, matrice active (texte affiché et non plus reflété) etc.

qu'un élément libérateur. Nous avons été confrontés à ce problème dans ForSIC, et nous l'avons résolu en déclinant l'ontologie générique de compétences sous une forme plus explicite, c'est-à-dire moins contraignante parce que plus transparente pour l'utilisateur. Cette méthodologie qui consiste à ramener en tâche de fond les ontologies posant des problèmes d'usage ou de compréhension, tout en conservant leur potentiel de formalisation, nous semble être une solution au problème de résistance de différents champs.

Un autre problème récurrent dans l'élaboration d'ontologies pour l'aide à la résolution de problèmes (la formation des usagers étant, en soi, un problème à résoudre) est que l'ensemble des représentations du savoir destinées à la résolution de problèmes sont fortement et nécessairement affectées par la nature du problème à résoudre et par les stratégies d'inférence qui peuvent lui être appliquées. Là encore, l'orientation méthodologique choisie dans ForSIC qui consiste à faire coopérer des types ontologiques distincts, dans une architecture hypertextuelle permettant de les relier et de les enrichir mutuellement, nous semble être une solution pérenne et portable : il n'est dès lors plus nécessaire de se mettre en quête de ces ontologies « top level » (de haut niveau) censées permettre de rassembler sous une ontologie unique, l'ensemble des domaines de connaissance<sup>97</sup>.

Enfin, le dernier problème posé par les ontologies est celui de leur mise en œuvre au niveau des liens hypertextuels qui permettent de naviguer dans les concepts ou notions qu'elles articulent :

*« Quelques systèmes hypermédia conceptuels exposent leur ontologie et la rendent explicitement navigable ; d'autres rendent le schéma de classification plus implicite. L'ontologie doit-elle être visible durant l'activité de liaison ? Quand elle devient un modèle sophistiqué de rôles et d'axiomes plutôt qu'une simple arborescence statique, choisir un concept devient un exercice d'expression d'une pensée construite. »* [Carr et al. 01]

Nous pensons, en accord avec l'argumentaire de Carr, qu'il est important de cantonner les ontologies à un niveau implicite d'interfaçage, laissant en revanche à un niveau explicite les thésaurus ou les vocabulaires sur lesquels elles reposent (comme c'est le cas dans ForSIC pour le modèle de domaine, directement accessible via l'outil auteur SABRE). Le choix inverse risquerait d'amplifier et de poser de nouveaux problèmes de navigation, d'orientation et de compréhension. Les ontologies qui mettent en œuvre une axiomatique élaborée doivent être utilisées de manière transparente (dans la création ou dans la recherche) et limitées à ces aspects (création et recherche d'information via des outils dédiés). Pour les aspects centrés sur l'utilisateur, il faut être capable de proposer une ontologie interface, c'est-à-dire une catégorie particulière d'ontologie, réduite à une série de primitives (matrice générique), qui peut elle-même renvoyer, à chacun de ces niveaux, vers des représentations explicites (termes du modèle de domaine, compétences explicites).

L'outil auteur SABRE qui entre dans sa phase finale de test et doit être ouvert à l'ensemble du collectif nous paraît à ce titre exemplaire<sup>98</sup> : il offre dans une interface unique la possibilité de disposer des

<sup>97</sup> ontologies qui restent à ce jour – et probablement pour longtemps encore – un improbable graal de l'ingénierie des connaissances et du web sémantique.

<sup>98</sup> pour le descriptif et les spécifications complètes de cet outil on se référera à [Al-Tawki & Tazi 01] et à la thèse de Al-Tawki Yahya, *Applications des systèmes multiagents aux environnements de création de support de cours multimédia*, Université de Toulouse 1, sous la direction de M.F. Barthet, 2001.

## 7.5. Bilan et perspectives.

Nous présentons ici quelques éléments de conclusion sur ce projet et revenons sur les limites que posent certains de ses aspects, dont l'utilisation des ontologies.

Concernant sa finalité en tant que dispositif éducatif de formation, FoRSIC permet d'atteindre les buts fixés dès 1976 par [Illich 76 p.81] qui, en quête de nouveaux modèles éducatifs, distinguait quatre types de réseaux pouvant – devant – être mis en œuvre pour l'ensemble des ressources éducatives (« *educational resources* ») :

- « *Services de référence [pour les ressources matérielles] (...) pour faciliter l'accès à des choses ou à des procédés entrant dans le cadre d'un apprentissage formel (...)* ;
- *échanges de compétences – pour permettre aux gens de lister leurs compétences, les conditions auxquelles ils veulent qu'elles puissent servir de modèle à ceux qui veulent les acquérir, et l'adresse à laquelle ils peuvent être contactés ;*
- *rencontre de pairs – un réseau de communication permettant aux gens de décrire l'activité d'apprentissage dans laquelle ils souhaitent s'engager, dans l'espoir de trouver un partenaire (...)* ;
- *services de référence [pour les personnes] – qui peuvent être listées dans un répertoire fournissant les adresses et des descriptions effectuées par les gens eux-mêmes des professionnels, des para-professionnels et des indépendants ainsi que les conditions d'accès aux services qu'ils proposent. »*

FoRSIC permet effectivement de déployer ces différents niveaux d'entrée : le premier type de réseau dépeint par Illich correspond à la modélisation des actes documentaires dans leur acception la plus large, permettant l'accès à différents types de ressources présentes dans le dispositif. Le second type de réseau est explicitement présent au travers de la modélisation des compétences (« *échanges de compétences* »). Le troisième témoigne de la volonté de structurer le collectif autour d'un réseau de pairs (formateurs à la recherche d'information). Le quatrième enfin rassemble l'ensemble des informations administratives, personnelles et concernant l'environnement de formation (personnes-ressource) de chaque membre du collectif.

A ce titre il est à notre connaissance l'un des rares dispositifs à proposer, à ce niveau d'aboutissement, ce type d'architecture.

Pour autant, FoRSIC pose également un certain nombre de questions qui restent ouvertes. Elles sont essentiellement liées à l'utilisation des ontologies et à la difficulté d'appliquer certains des formalismes ontologiques opérants dans le champ de l'ingénierie des connaissances<sup>96</sup>, à celui de la connaissance déployée sur les réseaux – web sémantique notamment. Dans ce dernier cadre, les problèmes liés aux ontologies résident principalement dans le nombre d'applications permettant de les utiliser : la construction d'ontologies est habituellement considérée comme coûteuse en temps aussi bien qu'en investissement.

Un problème plus essentiel est lié au fait que beaucoup de domaines de connaissance résistent à une formalisation précise : il est aisé, pour certains d'entre eux de déterminer des points en deçà ou au-delà desquels le potentiel de formalisation et de partage offert par les ontologies devient davantage une limitation

<sup>96</sup> qui s'intéresse principalement à des environnements centrés, orientés ou limités en terme de corpus ou de notions.

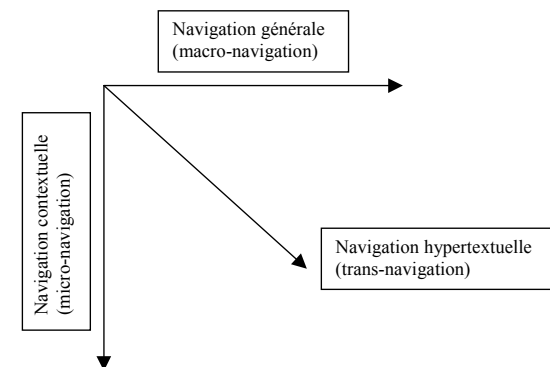


Fig. 14 : Principe de double orientation.

Les deux axes horizontaux et verticaux peuvent être intervertis selon différents critères : l'utilisation la plus courante associe l'axe horizontal aux sections et sous-sections principales (d'un site ou d'un hypertexte), et réserve l'axe vertical (le plus étendu) pour le niveau le plus riche (possibilité d'ascenseur – scrolling). Précisons ici que ce principe de double orientation relève bel et bien de ce niveau que nous qualifions de sémantico-temporel, et ce même s'il semble reposer sur des principes davantage liés à la spatialité dans la manière dont sont disposées les informations. Qu'il s'agisse de la macro ou de la micro-navigation, les axes autour desquels elles gravitent sont bien des axes temporels et non spatiaux. Au travers de l'activation des liens qui composent ces axes, c'est toute la linéarité (la linéarité n'excluant pas la récursivité) de la navigation qui se donne à lire.

Les tables des matières, les sommaires, certains index ou thésaurus, tous ces éléments ont comme ambition commune d'offrir au lecteur-utilisateur des vues synoptiques. C'est le niveau de granularité dans la focalisation offerte par ces vues qui déterminera à la fois le maintien ou la perte d'orientation et la capacité ou l'incapacité à construire, à partir de ces repères d'orientation, des modèles mentaux appropriés pour traiter l'information affichée. On retrouve ces phénomènes évoqués d'une manière plus générique dans la littérature sous le nom de « structures arborescentes ». Si leur utilisation est désormais un incontournable de la réalisation d'hypertextes (avec toutes les libertés et les variantes<sup>243</sup> que supposent le choix des nœuds de l'arbre, la densité de ramification choisie, etc.), leur utilisation exclusive, tout comme celle des liens « enrichis » évoqués plus haut, ne suffit à offrir l'ensemble des repères nécessaires à une navigation optimale du point de vue de l'utilisateur :

<sup>243</sup> on a par exemple pour habitude de distinguer entre les structures arborescentes orientées-objet (art ancien → égyptien → grec → romain) et celles orientées-tâche (acheter une voiture → consulter le prix → contacter un vendeur).

« ossature hiérarchique ou structure arborescente : aide l'utilisateur à développer un modèle mental du système et à maintenir son orientation.. (...) Mais un utilisateur n'a aucun moyen de savoir si les liens impliquent une relation d'une partie vers un tout, ou si les nœuds suivent une logique séquentielle s'étendant en largeur ou en profondeur, ou si l'information en bout des branches de l'arborescence soutient ou complique les nœuds internes. » [Young 90 p.240]

### 7.5.2.2. Sémantique externe (niveau spatio-sémantique).

On ne saurait parler de navigation hypertextuelle sans prendre en compte un certain nombre de facteurs liés à la dimension « spatiale » de cette activité. Ce sont ces critères que nous voulons maintenant analyser sous l'angle de la sémantique « externe » qui permet de gérer l'échelle des rapports possibles entre les différents espaces « fenêtrés » d'une navigation et leur influence en termes de feedback sur les contenus affichés et sur les représentations mentales qu'ils contribuent à forger chez l'utilisateur.

Depuis sa création, le web est fondé sur la « page-écran » comme unité atomique d'information. Jusqu'à il y a peu, les problèmes liés à la navigation étaient pensés dans ce contexte où l'unité de navigation (c'est-à-dire en fait l'unité d'adressage) était la même que l'unité de vision : 1 page = 1 adresse = 1 unité de navigation. L'apparition des techniques de cadres (frame)<sup>244</sup>, en venant remettre en cause le design originel du web, ont également conduit à repenser les problèmes liés à la navigation dans une nouvelle optique.

Le principe des frames est simple : pour une page-écran donnée, une « page-mère » (une adresse unique donc) peut contenir autant de « pages-filles » (et d'adresses différentes) que souhaité par l'auteur. A la question traditionnelle de savoir où se trouve l'information sur Internet (par exemple la page personnelle de Hans Dietrich), les plus pragmatiques répondaient : « à l'adresse suivante : <http://www.univ-stuttgart.edu/> ... », allant parfois jusqu'à expliquer que derrière cette adresse se trouvait en fait un numéro IP de machine (193.49.53.1 par exemple), laquelle machine était un serveur installé dans les locaux de l'université de Stuttgart, en Allemagne. Avec l'arrivée des frames et l'engouement qu'elles suscitent chez les auteurs d'hypertexte<sup>245</sup>, le problème prend un nouveau tour : si l'information affichée via l'interface de navigation conserve un adressage géographique, il devient plus compliqué de l'identifier avec précision, puisque derrière l'unité de vision peuvent se dissimuler plusieurs unités de navigation (autant que de frames). S'il est vrai que le codage HTML qui permet de créer un système de frames est simple (trois balises), il en va tout autrement de celui permettant de gérer les liens entre ces frames. De nombreux sites web sont encore présents sur la toile dans lesquels le résultat d'un clic confine à l'aléatoire absolu et amplifie de manière démesurée les problèmes de désorientation. Il est en effet possible, dans un système de frames de commander l'ouverture d'un nœud depuis une ancre :

- dans l'une des autres zones composant le système de frames (autre que celle où se situe l'ancre) ;

<sup>244</sup> 1998 pour leur utilisation « grand public ». (source <http://www.w3c.org>)

<sup>245</sup> les principaux avantages des frames (outre la relative simplicité de leur codage HTML) sont qu'elles permettent de pérenniser les niveaux contextuels de navigation offerts par les structures arborescentes, en laissant par exemple le sommaire constamment affiché à l'écran. Ajoutons également qu'elles offrent de nombreuses possibilités combinées à l'utilisation du JavaScript.

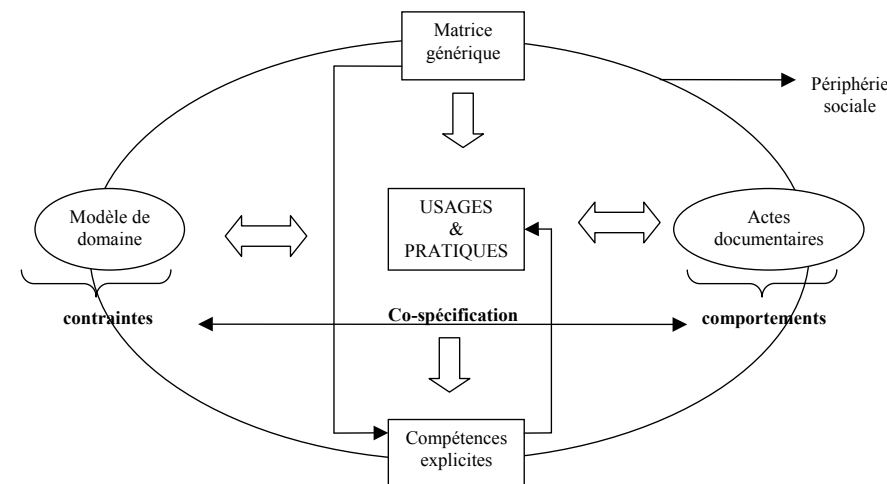


Fig. 23 : FoRSIC, une organisation hypertextuelle.

### 7.4.4.3. Un modèle pour une pragmatique de la connaissance ?

« En somme il y a autant d'ontologies qu'il y a de sciences : chacune, par diverses procédures ou rituels d'objectivation, crée et recrée son objet, établissant peu à peu son domaine d'objectivité. (...) Bref, faire de la science, ce n'est pas vérifier une ontologie, c'est se priver d'ontologie préconçue, et cette privation me paraît constitutive, dans la mesure où elle permet de constituer les objets scientifiques en passant outre à l'objectivisme de sens commun. » [Rastier 95]

L'architecture finale issue des croisements de ces différents niveaux et reproduite ci-dessus peut être posée comme modèle référent pour la mise en place d'une pragmatique de la connaissance dans une organisation de type hypertextuelle.

D'abord parce qu'elle renforce la pertinence individuelle de chacune des entrées la composant.

Ensuite parce qu'au vu des croisements identifiés et des interactions permises, elle permet leur enrichissement mutuel et dynamique.

Parce qu'elle est portable – sous couvert de certains ajustements – indépendamment du champ dont il est ici question.

Parce qu'enfin, elle est en adéquation avec les principes régissant le paradigme de la coopération tel que défini plus haut (cognition distribuée, LPP, conversion des connaissances, indexicalité, etc ...)

différents, et permet au système dans son entier d'évoluer de manière non-linéaire mais non entièrement entropique.

Afin d'illustrer ce propos, revenons sur quelques-uns des couplages les plus significatifs au sein de cette organisation :

- matrice générique et compétences explicites sont explicitement liées dans l'usage<sup>94</sup>, l'une permettant la production des autres, et ces dernières pouvant être amenées à infléchir ou à modifier l'orientation de la première (dans la phase de stabilisation du projet, nous avons ainsi opéré quelques substantielles modifications dans les termes retenus pour figurer dans la matrice générique) ;
- le modèle de domaine est lié aux compétences explicites à deux niveaux d'échelle, complémentaires :
  - celui du vocabulaire : celui qui est présent dans le modèle de domaine est repris tel quel dans les intitulés de compétences explicites, et vice-versa ;
  - celui des niveaux de formation (culture, expérience, stratégie) : il est possible de lire le modèle de domaine à l'aune des ces trois niveaux selon deux modes, deux directions :
    - quand on le parcourt horizontalement, plus on « avance », dans l'arborescence sur laquelle il est structuré, plus on tombe, dans les derniers niveaux, sur des termes témoignant d'un niveau de connaissance « stratégique » ;
    - quand on le parcourt verticalement, la disposition en premier lieu des « types d'outils », puis des « techniques d'interrogation » et enfin des « scénarios d'usage », rend elle aussi compte d'une progression vers des connaissances et des niveaux de maîtrise de plus en plus « experts »<sup>95</sup> ;
- les actes documentaires sont liés :
  - à la matrice générique, ce qui permet par exemple de contextualiser différemment les termes présents dans la classe « Action »,
  - aux compétences explicites, ce qui permet d'opérer un tri et des recoupements qui correspondent à chacun des actes définis,
  - au modèle de domaine, dont ils reflètent l'organisation en trois parties distinctes (« identifier » renvoie à la partie sur les « types d'outils », « croiser, associer » à celle sur les « procédures d'interrogation », « analyser » et « exploiter » à celle sur les « scénarios d'exploitation »).

<sup>94</sup> celui de l'interface de FoRSIC mais également celui des usages documentaires.

<sup>95</sup> il s'agit de considérations globales. Le « bon » parcours de lecture permettant de valider la pertinence de ce modèle de domaine n'est ni vertical, ni horizontal mais bien transversal, puisque que c'est dans cette transversalité que se donne à lire la nature de la recherche d'information comme processus d'apprentissage et non dans l'une ou l'autre des linéarités précédentes.

- dans la même zone que celle où se situe l'ancre ;
- en lieu et place du système de frame (dans la même session de navigation) ;
- dans une autre fenêtre de l'interface de navigation utilisée (ce qui revient à ouvrir une nouvelle session de navigation)

Pour mémoire, nous citerons également les problèmes qui n'entrent pas directement dans le champ de ce travail mais qui demeurent liés à l'utilisation de ces frames comme l'impression des pages « framées » et leur référencement délicat voire impossible tant pour les moteurs de recherche que pour les signets des utilisateurs.

### **7.5.3. Etablir des cartes d'orientation.**

Il s'agit de la solution la plus utilisée et de celle qui a fait l'objet du plus grand nombre d'études critiques soulignant les problèmes spécifiques de ce type de solution. [Balasubramanian 94] mentionne quatre grandes catégories d'outils :

- navigateurs graphiques,
- vues différentes,
- cartes et diagrammes synoptiques,
- cartes conceptuelles.

Quelle que soit la catégorie choisie, l'avantage de ces solutions est double : d'une part permettre à l'utilisateur, par des vues globales, de développer des stratégies cognitives adéquates, et d'autre part de préserver son sens de l'orientation. Cependant, si ces solutions s'avèrent parfaitement efficace pour des hypertextes à structure simple, l'augmentation du volume d'information et de la complexité des structures permettant de s'y repérer, augmente du même coup le niveau d'échelle des cartes permettant d'en rendre compte, la carte risquant de devenir illisible pour l'utilisateur n'en maîtrisant pas les codes<sup>246</sup>.

On en arrive parfois au résultat inverse et des études expérimentales ont permis de montrer les lacunes de ces approches. Pour les cartes de type « graphique » mettant en évidence une topologie structurelle de l'espace informationnel :

*« Il n'y a en général pas de meilleurs résultats en apprentissage ou d'orientation dans l'hyperdocument quand le système dispose d'une « carte des contenus » que quand il n'en dispose pas. Dillon et ses collègues (1993) ont expliqué ce résultat (maintes fois répliqué) en disant que l'espace de navigation n'est pas un espace physique dans lequel on s'oriente, mais un espace sémantique dans lequel on traite des contenus. » [Tricot et al. 98]*

Une logique d'opposition se met alors en place : *« Le fait qu'une vue globale structurelle permette de rendre compte d'une structure globale est apparemment atténué par le fait qu'elle entraîne l'attention du lecteur loin de la structure locale de l'hypertexte. » [Hofman & Oostendorp 98 p.123]*

<sup>246</sup> ces arguments sont notamment développés par [Young 90 pp.238-240].



Quant aux cartes conceptuelles :

« (...) *Ces résultats* [de l'étude menée dans cet article] *suggèrent qu'une carte conceptuelle peut gêner la compréhension des lecteurs novices dans le domaine considéré, parce qu'elle attire leur attention sur la macrostructure textuelle au détriment de l'attention portée à la microstructure.* » [Hofman & Oostendorp 98 p.119]

#### **7.5.4. Lutter par l'établissement de parcours de navigation.**

L'utilisation de parcours privilégiés ou guidés de navigation<sup>247</sup> est souvent utilisée en complémentarité avec les cartes de navigation, pour augurer l'utilisateur vers un itinéraire idéalisé permettant d'accroître sa compréhension de l'hypertexte ainsi parcouru ou (pour les hypertextes éducatifs de type tutoriel) de s'assurer qu'il accède en temps voulu à des nœuds faisant office de pré-requis. [Balasubramanian 94] distingue ici deux catégories<sup>248</sup> :

- Chemins et sentiers (« *Paths and Trails* »)
- Tours guidés (« *Guided Tours* »)

Ces deux solutions font l'objet de nombreuses recherches<sup>249</sup> au vu de leur pertinence pour l'ensemble des environnements d'apprentissage, de l'importance croissante de ces derniers sur le web, et des réflexions et expérimentations pratiques dont elles sont l'objet depuis des années en psychologie cognitive et comportementale<sup>250</sup>.

#### **7.6. Conclusion et pistes de réflexion.**

Comme nous l'avons souligné en introduction à cette partie, la difficulté principale concernant les stratégies de navigation provient d'abord des différents niveaux de maîtrise de l'utilisateur (naviguant) et du concepteur (mise en place des possibilités de parcours). [Rhéaume 93] : « (...) *ces problèmes [désorientation et surcharge cognitive] deviennent des qualités si cette flexibilité entre les zones d'information facilite la génération et l'organisation des idées, l'établissement des lieux de résolution de problèmes et même l'écriture en collaboration.* »

Nous proposons maintenant de dégager des axes de réflexion plus génériques que les solutions existantes retracées dans cette partie et qui nous amèneront à aborder notre dernier point : l'établissement et la distinction de deux invariants de la navigation hypertextuelle : le « browsing » et le « searching ».

<sup>247</sup> parfois appelés PPI (Parcours personnalisés d'information) ou PPA (parcours personnalisés d'apprentissage).

<sup>248</sup> « *Un jeu de liens s'appelle une navigation si l'objectif recherché est précis, un tour guidé si le cheminement est proposé par un tuteur et un broutage ou butinage si le lecteur évalue chaque îlot d'information à son mérite.* » [Rhéaume 93]

<sup>249</sup> pour un état de l'art sur cette question précise on renverra le lecteur à l'article de [Carr et al. 99b] qui revient sur les différences de nature et de fonction entre « *tours* » et « *trails* » et sur leurs implications quand ils sont distinctement ou simultanément mis en œuvre. « *Alors que les visites guidées [tours] servent à introduire un sujet, les sentiers [trails] se prêtent à des applications plus avancées.* »

<sup>250</sup> on consultera pour ce type de référence l'article de [Weinreich & Lamersdorf 00] qui en plus d'exemples détaillés d'applications dédiées, comprend un état de l'art des études menées en psychologie cognitive et en psychologie comportementale pour évaluer les stratégies de navigation et les types de surcharge liées.

Nous avons alors entrepris un nouveau travail de filtrage et de reformulation pour « traduire » ces compétences génériques en compétences spécifiques, c'est-à-dire rendant compte de pratiques effectives de formation. A l'issue de cette nouvelle étape, près de 300 compétences explicites ont été validées, toujours en concertation avec le collectif. C'est ce référentiel de 300 compétences explicites qui est présent dans l'interface et sur lequel le collectif se positionne. La possibilité d'utiliser les classes comme autant d'entrées d'index demeure, mais n'est utilisée que par les formateurs maîtrisant le mieux l'interface, les autres préférant une navigation plus classique, de balayage du référentiel.

En opérant cette reformulation, nous avons de fait renforcé la nature ontologique de la matrice initiale : une même combinaison de termes peut désormais donner lieu à la formulation de plusieurs compétences explicites. Ainsi la combinaison : « *enseigner* (action), *dans un contexte informationnel général* (Environnement), *pour des étudiants* (Bénéficiaire), *une méthode, une technique* (Délivable) » donne lieu à près de dix-neuf compétence explicites : connaître les langages documentaires, vérifier une référence bibliographique, construire une équation de recherche, interroger les moteurs de recherche, interroger les banques de données, etc.

#### **7.4.4. Couplages structurels.**

##### **7.4.4.1. Une ingénierie de l'usage.**

« *Pour limiter l'explosion hypertextuelle, l'ingénierie de la connaissance s'attachera à déterminer les espaces de représentation centraux d'un domaine de connaissance, en fonction des objectifs et des intérêts des utilisateurs.* » [Lévy 91 p.57]

Depuis le moment où l'idée du projet de recherche FoRSIC a été formulée, jusqu'à celui de son ouverture aux utilisateurs, l'idée de déterminer une ingénierie de l'usage a toujours été présente et a représenté l'axe moteur du projet. Cependant, quelque soit le champ choisi, et tout particulièrement pour celui relativement peu formalisé en dépit des apparences, de la formation à la recherche documentaire, ces usages sont évidemment différenciés et hétérogènes. Ils relèvent de différents niveaux de compétence et de pratique, et mettent en œuvre des styles cognitifs tout aussi complexes et hétérogènes.

C'est pour pallier cette hétérogénéité, pour identifier des invariants, et pour faciliter une dynamique et une ré-ingénierie des pratiques dont ces usages rendent compte, que nous avons relié l'ensemble des modélisations et des niveaux ontologiques jusqu'ici évoqués au sein d'une organisation de type hypertextuelle que nous présentons maintenant.

##### **7.4.4.2. Une organisation hypertextuelle.**

L'architecture d'ensemble de ces éléments est bien de nature hypertextuelle : chaque entité fonctionne comme une entrée possible, toutes sont liées, et chaque évolution de l'une d'entre elles (nouveaux termes dans le modèle de domaine, nouvelles compétences explicites, modification de la matrice générique ...) a des répercussions directes sur l'ensemble des autres, mais à des niveaux d'échelle et de perception



– et elle peut être « ce que l'on sait expliciter » – il s'agit alors de trouver les moyens de partager, toujours à l'échelle du collectif, ce niveau « subjectif » d'explicite. Dans l'un comme dans l'autre cas, le recours à une ontologie s'avèrera nécessaire.

Développée avec l'outil SEE-K comme le lien entre des individus, des connaissances et des besoins, nous avons utilisé la méthodologie d'élaboration d'une matrice de compétence, dénommée A.B.E.D. (Action Bénéficiaire, Environnement, Délivrable) proposée par la société Trivium. Il s'agit en fait d'une ontologie orientée-tâche organisée autour de quatre dimensions (classes) qui reprennent les quatre questions fondamentales dans le cadre de la résolution de problèmes :

- classe (dimension) Action : que suis-je capable de faire ?
- classe (dimension) Bénéficiaire : à qui cela bénéficie-t-il ?
- classe (dimension) Environnement : de quoi ai-je besoin pour le faire ?
- classe (dimension) Délivrable : au final, quel élément nouveau cela apporte-t-il ? (quel type de ressource, quel type de savoir-faire, etc ...)

#### **7.4.3.1. Compétences génériques.**

Nous avons alors, sur la base de ces quatre dimensions, sélectionné pour chacune d'entre elles une série de cinq termes (concepts), en respectant les règles méthodologiques indiquées par [Gruber 93], permettant ainsi la production automatique par croisement d'un référentiel de 625 compétences génériques (5 x 5 x 5 x 5). Nous avons ensuite opéré un tri « manuel » dans le référentiel produit – en concertation avec les membres du collectif – pour éliminer celles ne faisant pas sens. Sur la base des 625 produites, un peu moins de 400 ont ainsi été validées.

Action	Bénéficiaire	Environnement	Délivrable
Identifier, chercher	soi-même	Contexte informationnel général	approche théorique
Concevoir, étudier	étudiants, apprenants	Champ disciplinaire	Méthodes et techniques
Réaliser, produire	Collaborateurs, Formateurs	Techniques infométriques	Valeur ajoutée à l'information
Enseigner, informer	Institution, université	Réseau humain	Expertise, conseil
Valider maintenir	Partenaires, fournisseurs	conditions de réalisation (aspects matériels)	norme, diplôme.

**Tableau 14 : matrice ontologique de compétences génériques.**

#### **7.4.3.2. Compétences explicites.**

A l'usage, et bien que l'on puisse accéder à ces compétences dans l'interface SEE-K en utilisant les noms des classes (dimensions) comme autant d'entrées d'index, le collectif a fait part de la difficulté qu'il éprouvait à se positionner de manière discriminante sur des chaînes de caractère comme : « *concevoir, étudier, pour des partenaires, des fournisseurs, selon des conditions de réalisation, des méthodes, des techniques.* »

### **7.6.1. Pour une pédagogie « moderne » de l'hypertexte.**

Nombre d'universités anglo-saxonnes proposent depuis de nombreuses années des cours d'écriture multimédia prenant en compte les particularités de l'hypertexte. Il ne paraît actuellement ni possible ni raisonnable de faire l'économie de ce genre de formation avant de vouloir ou de pouvoir proposer des processus d'optimisation de la navigation valables à l'échelle du réseau global. Pour être cohérente, de la même manière que les travaux sur l'apprentissage de la lecture ont depuis longtemps établi des jalons correspondant à différents stades de développement cognitifs, la lecture de l'hypertexte repose également sur différents modes ou styles cognitifs partagés pour la plupart par ceux mis en place dans la lecture traditionnelle et d'autres, plus spécifiques qui nous paraissent déterminants<sup>251</sup> et pour lesquels les travaux recensés par [Besiat 93]<sup>252</sup> pourraient servir de base en établissant des recoupements avec les différentes théories de l'apprentissage et les styles cognitifs associés<sup>253</sup>.

La plupart de ces approches considèrent la navigation d'abord comme une activité relevant d'une « gestion de tâche ». Différents types d'aides peuvent alors être proposées<sup>254</sup>. [Fraissé 97 p.254] propose de distinguer un « *modèle de tâche* » (MT), un « *modèle de navigation abstraite* » (MNA) et un « *modèle de navigation effective* » (MNE) interdépendants permettant de finaliser et de mettre en œuvre des scénarios de navigation opératoires. [Sabah & Zock 92 p.12] se placent quant à eux dans le cadre de la théorie des graphes pour proposer la notion d'AFT (« *Arbres de foyers thématiques* ») rendant compte du fait « *qu'un lecteur attend des informations différentes en différents points de la chaîne discursive* ».

### **7.6.2. Le recours à l'antique.**

Comme ce fut le cas pour les parties consacrées à la rhétorique des figures de l'hypertexte ou au rôle central des activités mnésiques et à leur utilisation dans une rhétorique du discours, il nous semble qu'en termes cette fois plus « communicationnels », pourraient être repris et adaptés les codes établis depuis l'antiquité ayant trait par exemple à l'art de la persuasion<sup>255</sup>.

<sup>251</sup> voir notamment le point 6.3.7. « Fonction(nalités). » de ce chapitre.

<sup>252</sup> ces travaux permettent notamment d'établir trois catégories d'activités méta-cognitives pouvant avoir une influence sur la performance en lecture : la connaissance des stratégies de compréhension des textes, la connaissance des structures des textes et la gestion de la compréhension. [Besiat 93 p.135]

<sup>253</sup> <http://tip.psychology.org>

<sup>254</sup> « aides à la gestion de la tâche : il s'agit essentiellement d'aider le sujet à se représenter de façon opérationnelle le but qu'il poursuit ; dans certains cas il faudra aider le sujet à faire évoluer cette représentation, tandis que dans d'autres, il faut l'aider à maintenir cette représentation stable (on sait que le problème de désorientation dans l'hyperdocument est souvent dû à un problème de maintien en MDT [modèle de tâche] de la représentation du but). » [Tricol et al. 98 p.11]

<sup>255</sup> [Butera 02] rappelle le rôle fondamental dans ce cadre de « l'effet de primauté » d'une part (« quand deux messages sont émis l'un à la suite de l'autre, l'oubli est similaire pour les deux et le premier message a plus d'impact. » ) et de « l'effet de récence » d'autre part (« quand un long délai sépare les messages, l'effet du premier s'est estompé, alors le message le plus récent exerce le plus fort effet de persuasion. »)

### 7.6.3. Navigation tangentielle.

C'est [Winkin 81 p.243] qui pose et étudie ce phénomène dans le cadre des structures de la communication : « Si la déclaration (a) par la personne A est suivie de la déclaration (b) par la personne B, et si (b), d'une part, reconnaît la volonté de communication de A, mais, d'autre part, néglige à la fois le contenu de (a) -le message- et le but de A -son émetteur-, nous obtenons une réponse tangentielle. » Par nature, l'hypertexte – c'est-à-dire l'ensemble des possibilités de liaison qu'il autorise et l'ensemble des stratégies de navigation associées – est de nature tangentielle. [Balasubramanian 94] reprend et évoque cet état de fait qu'il décrit ainsi : « L'hypertexte offre comme capacité inhérente la création de classifications latérales. » On parle de « latéralité » en recherche documentaire à propos de la reformulation de requêtes. La « création de classifications latérales » dont il est ici question désigne tout à la fois le basculement, via la navigation, vers des catégories d'hypertextes différentes, et le recours à des styles cognitifs différents que chaque usager met en œuvre dans son parcours de navigation<sup>256</sup>. Il s'agit bien là d'un nouvel invariant de l'organisation hypertextuelle<sup>257</sup>.

### 7.6.4. Invariants procéduraux et déclaratifs.

Sous sa forme actuelle, l'organisation hypertextuelle mise en place sur le web est un ensemble de liens procéduraux<sup>258</sup> au service de connaissances déclaratives (essentiellement textuelles)<sup>259</sup>. Or l'ensemble des interactions entrant en jeu dans le cadre d'une session de navigation reposent sur les choix d'activation – ou de non-activation – de ces liens et sur les indications qu'ils fournissent. Toute navigation est donc procédurale, et – du fait des lacunes propres aux liens (auxquelles nos invariants doivent permettre de répondre) – entraîne « nécessairement » une surcharge cognitive ou des problèmes de désorientation.

Pour résoudre ces problèmes, l'objectif à atteindre est celui d'une coopération idéale entre l'utilisateur responsable de la navigation d'une part (qu'il soit « auteur » ou « lecteur ») et les liens et les connaissances liées d'autre part. [Gleizes & Glize 99] définissent ainsi les conditions d'une coopération idéale :

*« Les caractéristiques de la coopération idéale (...) sont aussi celles d'une coopération totale où la moindre activité est bénéfique pour autrui :*

<sup>256</sup> voir aussi le point 7.3. déjà développé sur le « syndrome d'Elpénor ».

<sup>257</sup> sur lequel nous reviendrons dans le point 7.7 « Browsing VS searching » en montrant comment et sur quoi s'opère la tangentialisation dans l'un et l'autre cas.

<sup>258</sup> [Hofstadter 85 p.168] « Le plus souvent, lorsque l'on a besoin d'une procédure, on veut qu'elle puisse varier en fonction du contexte [ce qui n'est pas le cas des liens actuels]. Il faut donc que cette procédure ait un moyen de voir ce qui est stocké en mémoire et de choisir ses actions en fonction de ce contenu, ou qu'on lui fournisse explicitement une liste de paramètres la guidant dans le choix des opérations à effectuer. » L'enjeu est ici d'extérioriser l'information que contiennent ces procédures, de la rendre disponible à l'utilisateur au moins autant qu'au seul système.

<sup>259</sup> [Hofstadter 85 p.407] « Une connaissance est dite déclarative si elle est stockée explicitement. Dans ce cas, le programmeur et le programme peuvent la « lire » comme si elle se trouvait dans une encyclopédie ou dans un annuaire. (...) La connaissance procédurale, à l'inverse, n'a pas de forme explicite. Elle est codée exclusivement sous forme de programmes. »

Le déploiement spiralaire de ces différentes étapes du processus de conversion de connaissance, fait qu'il peut s'appliquer – comme c'est visible dans le paragraphe précédent – à l'ensemble d'un processus de formation comprenant les trois niveaux, mais peut également servir à qualifier et à hiérarchiser le parcours propre à chacun de ces niveaux, pris indépendamment. En effet, et comme nous l'avons déjà rappelé, ce qui apparaît comme une étape finale dans la présentation synoptique de ces actes documentaires (tableau ci-dessous), n'est en fait qu'une étape à partir de laquelle un nouveau cycle de recherche est entamé, suivant les mêmes étapes, mais dans lequel l'utilisateur dispose alors d'un savoir « augmenté », et en tout cas différent de celui dont il disposait initialement.

Actes documentaires	Activités cognitives	Niveaux d'usage		
		Culture	Expérience	Stratégie
Identifier	Reconnaître (percevoir, traiter)	Comprendre les règles régissant l'organisation des différents outils	Comprendre l'organisation de l'information dans un champ disciplinaire	Mise en relation des modèles, des contextes, des situations.
Croiser, associer outils et demandes	Comprendre raisonner	Travailler sur des représentations formelles de l'information	Comprendre les situations, les différents contextes	Appréhender le niveau implicite de l'information : savoir le faire émerger.
Analyser (comparer, extraire)	Combiner	Trouver des réponses à des problèmes simples	Comprendre les pratiques, obtenir la représentation d'un champ disciplinaire.	Résoudre de problèmes non ou mal définis, dilemmes
Exploiter l'information obtenue	Produire	Synthèse de documents, résumé, bibliographie ...	Construction sociale, mise en évidence de relations.	Mettre en relation des éléments disparates, des événements et des réseaux, pour établir des projections dans le futur. Analyse multidimensionnelle

**Tableau 13 : La recherche d'information comme processus d'apprentissage.**

### 7.4.3. Matrice de compétences.

Tous les systèmes de recherche et de classement d'information fonctionnent sur la base de contraintes fortes (opérateurs booléens, langages documentaires ...), face auxquelles chaque usager adopte, selon son niveau de connaissance et de pratique, un comportement différent. Dans FoRSIC, nous nous sommes mis en quête d'un élément se prêtant à une catégorisation de type ontologique et capable de synthétiser et de caractériser simultanément des contraintes ET des comportements. Il fallait également que cet élément constitue un point de convergence entre des individus, des savoir-faire et des connaissances (indexicalité). Nous avons pour cela choisi la notion de compétence, définie par M. Authier comme « la forme que prend la connaissance chez un individu », ce qui nous permet de « réconcilier » deux approches traditionnellement antinomiques associées à cette notion : la compétence peut être une « connaissance fonctionnelle » – il s'agit de trouver les mots les plus adéquats pour en rendre compte à l'échelle du collectif

Les quatre activités cognitives retenues<sup>93</sup> sont : reconnaître (1), comprendre, raisonner (2), combiner (3), produire (4).

#### **7.4.2.3. Actes documentaires.**

En corrélant les trois niveaux de formation avec les activités cognitives nécessaires pour chacun d'entre eux, nous avons défini quatre fondamentaux (invariants) de la recherche d'information comme processus d'apprentissage :

- identifier : il s'agit d'identifier les transformations des données dans le système d'information pour permettre à l'utilisateur d'exprimer une requête claire et pertinente ;
- croiser, associer : il s'agit d'être capable de choisir l'outil de recherche le plus adapté au besoin d'information exprimé. Ce qui implique que l'utilisateur puisse disposer de catégories pertinentes auxquelles rattacher ces outils (moteurs, annuaires, bases de données, technologies agent, etc ...). Il peut alors croiser sa demande et les outils ainsi sélectionnés ;
- analyser : il s'agit d'être capable d'analyser l'information retrouvée afin de définir sa pertinence (en fonction de l'objectif et du cadre dans lequel elle sera réutilisée) ;
- exploiter : il s'agit enfin d'exploiter l'information dans une nouvelle production (itération), pour mettre en place une action déterminée, ou pour constituer un nouveau savoir.

Nous avons alors confirmé avec le collectif de formateurs du projet, et par un travail approfondi sur plusieurs ressources de formation, la validité des transversalités de ces différents modèles, tels que présentés dans le tableau synoptique ci-dessous (tableau 13).

Il faut ici repréciser que la construction de ces modèles et leur croisement ne s'est évidemment pas faite *ab abstracto*. Elle a été menée dans le cadre des principes retenus pour le projet FoRSIC et exposés plus haut. Nous avons notamment validé l'hypothèse de départ selon laquelle l'articulation de ces différents actes documentaires et des niveaux de formation auxquels ils renvoient, s'inscrit dans le modèle de conversion de connaissances décrit par [Takeuchi & Nonaka 95] : ainsi, pour la phase de socialisation qui concerne la connaissance tacite, le niveau de recherche d'information est peu ou mal formulé (niveau « culture »). L'étape suivante, l'extériorisation met en œuvre une structuration plus avancée, souvent articulée par la conceptualisation d'un domaine. La construction d'un savoir structuré autour de la recherche d'information se fait par combinaison de connaissances explicites qui vont être ordonnées, classifiées pour trouver une place logique dans un système de classification et dans un savoir hiérarchique (niveau « expérience »). Enfin, l'intériorisation (retour à un niveau de connaissances tacites) s'incarne dans une production particulière, dans un travail précis à réaliser (niveau « stratégie »).

<sup>93</sup> il s'agit d'une équivalence qui ne vaut que pour le champ de la formation à la recherche d'information : l'activité de définition permet aux apprenants d'acquérir des modes de raisonnement, celle de structuration correspond à la mise en œuvre de ces acquis, etc.

- *compréhension : un signal perçu doit être interprétable par un système coopératif. La compréhension mutuelle n'a pas à être postulée mais doit émerger de l'ajustement mutuel entre le système et son environnement.*
- *Raisonnement : toute information (tout signal interprété) doit avoir des conséquences logiques dans le système. En d'autres termes, une information doit apporter de la nouveauté.*
- *Action : les conclusions du processus de raisonnement doivent être utiles à l'environnement du système coopératif. »*

Sous leur forme actuelle, les liens procéduraux majoritairement en place sur le web impliquent à l'inverse :

- *« Incompréhension : un signal perçu est incompris ou possède de multiples interprétations (ambiguïté).*
- *Défaut de raisonnement : l'information reçue est déjà connue ou n'a aucune conséquence logique. (...)*
- *Inutilité de l'action : compte tenu de ses croyances courantes, la partie considère que la transformation de l'environnement qu'elle peut opérer n'est pas bénéfique à autrui. Cette situation englobe les notions de conflit et de concurrence. »*

Il paraît logique (et souhaitable) que la connaissance continue de son côté d'être stockée et affichée sous des formes déclaratives. L'optique inverse (connaissance procédurale) pouvant être assimilée à l'actuel web invisible (pages générées à la demande et/ou stockées dans des bases de données<sup>260</sup>).

C'est donc bien notre approche des liens qu'il faut modifier en les ramenant – nos invariants ont cette fonction – vers des modes aussi déclaratifs que possible. L'adéquation alors réalisée entre des liens déclaratifs pointant vers des connaissances également déclaratives, suffisant à garantir des résultats pour l'optimisation de l'ensemble des processus de navigation.

Le carré sémiotique ci-dessous résume et illustre notre propos.

<sup>260</sup> la plupart des bases de données fonctionnent sur un principe qui conditionne et subordonne l'accès aux connaissances déclaratives – articles scientifiques par exemple – à l'utilisation de logiques procédurales (langages de requête, logique booléenne ...). L'hypertexte n'est ni ne doit devenir une gigantesque base de données.

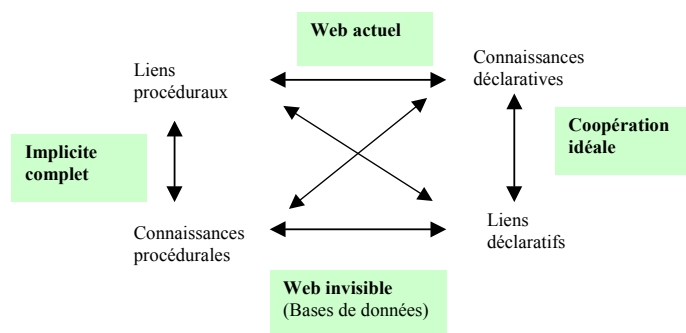


Fig. 15 : Carré sémiotique de la coopération idéale.

Dès lors, cette « coopération idéale » pourrait s'appliquer à l'ensemble des niveaux hiérarchiques proposés par [Cissé et al. 99] et couvrir tous les niveaux d'échelle et de granularité pouvant se présenter sur le réseau.

- « coopération : la coopération apparaît quand des actions individuelles contribuent aux actions des autres et vice-versa.
- collaboration : la collaboration est le fait de travailler ensemble dans l'exécution d'une certaine action, générant une compréhension commune et une connaissance partagée. Le résultat est ainsi imputable au groupe entier.
- co-décision : la co-décision concerne les décisions de groupe ou inspirées par le groupe, les acteurs étant soit indifférenciés, soit dotés de statut particulier. La crédibilité et la création de connaissances partagées et de reconnaissance mutuelle sont aussi importants que pour la collaboration. »

## 7.7. Browsing versus searching.

Si la richesse et la particularité de l'organisation hypertextuelle a comme clé de voûte les ancrs qui la composent, celles-ci sont à la fois facteur d'ordre et de chaos. Mais il ne s'agit pas d'une détermination *a priori* : rien n'indique, au moment où celles-ci sont posées si elles augmenteront l'entropie de l'ensemble de l'organisation ou permettront de la réduire (négentropie) ni à quel(s) niveau(x) d'échelle elles le permettront. Là encore – comme pour les genres hypertextuels – il s'agit d'une détermination *a posteriori*, qui est co-déterminée par l'usage et par la topologie à chaque instant renouvelée du réseau.

Nous avons choisi de terminer ce second chapitre de notre travail par le dernier et probablement le plus transversal des invariants jusqu'ici dégagés : la distinction et la caractérisation de deux invariants

une profondeur de sept niveaux d'arborescence. Chaque terme est accompagné d'une définition générique, d'une ressource pédagogique dédiée, d'un exercice ou d'un exemple dédié, et d'une ressource pédagogique plus générique permettant de recontextualiser le terme en question.

## 7.4.2. Scénarios d'usage.

Afin de pouvoir se reposer sur une vision cohérente et homogène de la recherche d'information comme processus d'apprentissage, nous avons proposé des modèles d'usage reposant sur des modèles d'apprentissage, isolant d'abord des niveaux de formation, des activités cognitives mobilisées pour ces tâches particulières, et quatre fondamentaux rendant compte des spécificités des pratiques de formation dans ce champ donné.

### 7.4.2.1. Niveaux de formation.

Les trois niveaux de formation distingués sont les suivants : culture, expérience, stratégie.

Le premier, culturel, concerne la connaissance que l'apprenant a des outils de recherche d'information et des usages qui leurs sont le plus fréquemment associés à un niveau général. Il s'agit d'une culture de base de l'information.

Le second (expérience), concerne les usages et les pratiques ayant cours dans un champ ou un domaine donné : nombre de champs (biologie, aéronautique ...) disposent en effet de fortes spécificités en termes d'outils, ou de langages documentaires permettant de les décrire.

Le troisième, stratégique, rend compte d'une manière différente d'utiliser l'information : il s'agit pour l'apprenant, de comprendre les aspects multidimensionnels d'une situation, afin de pouvoir se projeter dans le futur (analyse multidimensionnelle, techniques de veille, cartographies avancées, infométrie ...).

Notons ici que du fait de l'aspect itératif du processus d'apprentissage au cours de ces différents niveaux, chacun d'entre eux est, à chaque nouvelle itération, renforcé par l'expérience acquise, autorisant et favorisant un usage et des pratiques de plus en plus expertes, sur des corpus d'information de moins en moins explicites, de moins en moins « balisés ».

### 7.4.2.2. Activités cognitives.

Sur la base d'activités pédagogiques établies et reprises dans la plupart des environnements d'apprentissage, nous avons dégagé les activités cognitives dont elles rendent compte, afin de proposer une articulation spécifique au champ disciplinaire du projet (la formation à la recherche d'information).

Les quatre activités pédagogiques fondamentales retenues sont : définir (1), structurer, didactiser (2), présenter (3), évaluer (4).

compétences explicites constituent, au final, le vocabulaire partagé (ontologie) par l'ensemble des membres du collectif.

Nous disposerons alors, dans FoRSIC, d'une architecture ontologique à plusieurs couches :

- modèle de domaine,
- scénarios d'usage (actes documentaires + niveaux de formation + activités cognitives),
- matrice de compétences génériques,
- référentiel de compétences explicites.

Nous concluons par la validation de notre hypothèse de travail : il est possible de lier entre eux ces différents « niveaux » tout en veillant au respect des critères définis par Gruber. Ces liaisons ne sont possibles qu'à la condition de disposer comme cadre de mise en œuvre, d'une organisation de type hypertextuelle (telle que définie tout au long de ce travail). Nous indiquerons pourquoi l'architecture ainsi posée nous semble constituer un modèle pérenne pour la mise en œuvre d'une pragmatique de la connaissance.

#### **7.4.1. Modèle de domaine de la recherche d'information.**

##### **7.4.1.1. Du séquentiel au hiérarchique.**

Le champ de la recherche d'information ne dispose à notre connaissance d'aucun outil de représentation global, prenant en compte l'ensemble de ses formalismes et de ses spécificités, alors qu'il a paradoxalement pour vocation de servir à en élaborer pour divers domaines scientifiques. Nous nous sommes donc initialement lancés dans un balayage de ce champ afin d'isoler une liste de termes pertinents pouvant être utilisés dans le cadre d'une classification de type « index de notions ». Nous avons pour cela fait appel à différents outils linguistiques – dictionnaires en ligne, thesaurus WordNet, ASIS Thesaurus of Information Science, notamment<sup>92</sup> – ainsi qu'à l'expertise des membres du collectif de projet.

La deuxième étape a consisté à organiser ces termes sur un modèle de type thesaurus, passant ainsi d'une représentation séquentielle à une représentation hiérarchique.

##### **7.4.1.2. Du hiérarchique au contextuel.**

Le troisième temps de notre travail fut consacré à construire un réseau sémantique, sur la base des informations constituant le thesaurus, et à l'aide des spécifications de l'outil SABRE (point 6.2.2.).

##### **7.4.1.3. Du contextuel au multi-relationnel.**

Au final, nous avons élargi le spectre de ce modèle de domaine et parlons de multi-dimensionnalité dans la manière dont il permet d'articuler le reste des niveaux ontologiques présents dans FoRSIC et décrits plus bas. Il se compose actuellement d'une liste de près de cinq cent termes, déployée autour de trois entrées principales – types d'outils, procédures d'interrogation, scénarios d'exploitation – s'étendant chacune sur

navigationnels, celui du « browsing » et celui du « searching »<sup>261</sup>. Après avoir précisé le champ et le contexte dont ils sont issus, ainsi que le sens particulier que nous leur appliquons dans notre approche, ils nous permettront d'éclairer rétrospectivement la plupart des invariants précédents en y ajoutant un niveau de validation supplémentaire.

Ces deux notions sont rappelées par [Lucarella 90 p.81] dans un contexte qui est celui des stratégies de recherche d'information.

*« Cet article traite du problème de la recherche d'information dans un hypertexte. Dans ce contexte, le processus de recherche est envisagé comme un processus d'inférence qui peut être exécuté par l'utilisateur explorant le réseau hypertextuel (browsing), ou par le système, exploitant alors le réseau hypertextuel comme une base de connaissances (searching). » Et plus loin, p. 83 « Browsing : aller d'un endroit vers un objet [going from where to what] (en supposant que vous savez où vous vous trouvez dans la base de données et que vous voulez découvrir ce qu'elle contient à cet endroit). Searching : aller d'un objet vers un endroit [going from what to where] (en supposant que vous sachiez ce que vous cherchez et souhaitez trouver où cela se trouve dans la base de données). »<sup>262</sup>*

Nous reprenons l'idée selon laquelle le browsing relève d'une activité située, localisée, devant permettre un repérage de l'information (« from where to what »), s'opposant en cela au searching dont la nature tient d'abord au positionnement que le « chercheur » occupe dans une masse informationnelle au sein de laquelle il commence par s'inscrire avant de pouvoir s'y orienter (« from what to where »).

Nous choisissons par contre de ne pas retenir les distinctions opérées entre les utilisateurs et le système d'une part et le réseau hypertextuel et la base de connaissance d'autre part et allons maintenant expliquer pourquoi.

Il nous paraît erroné d'envisager les stratégies de recherche d'information dans un environnement hypertextuel (et partant celles de navigation), uniquement en fonction d'un objectif initial supposé constant tout au long du processus de recherche. En effet, si nous restons dans le cadre de la recherche d'information défini par [Lucarella 90]<sup>263</sup>, son raisonnement ne s'applique qu'au cas suivant, qui constitue l'une des toutes dernières étapes d'un processus classique de recherche d'information ou de recherche documentaire :

- on sait déjà ce que l'on cherche (nature de l'information, article de journal, publication scientifique ...), on a déjà identifié une série de mots-clés (nom de l'auteur, date de publication, mots du titre ...)
- et l'on sait où chercher (dans quels outils : bases de données, revues, encyclopédie ...)

<sup>261</sup> la littérature francophone sur la question propose une dichotomie Navigation / Exploration. Nous conservons ici les termes anglo-saxons, la « navigation » étant la nature des deux phénomènes étudiés

<sup>262</sup> ces deux notions sont pour la première fois posées par Brown P., « Linking and Searching within Hypertext », **Electronic Publishing**, vol.1 (1), pp.45-53, 1988.

<sup>263</sup> cette optique est partagée par la plupart des approches présentées en 4.5.1.1. Ainsi [Kleinberg 98 p.2] « Notre travail a comme origine le problème de la recherche sur le web, que nous pourrions définir grossièrement comme le processus permettant de découvrir des pages pertinentes en réponse à une requête donnée. »

<sup>92</sup> WordNet : <http://www.cogsci.princeton.edu/~wn>, ASIS Thesaurus : <http://www.asis.org/Publications/Thesaurus/isframe.htm>.

Or la nature du processus de recherche d'information sur Internet est radicalement différente<sup>264</sup> : l'ensemble de ces paramètres est défini simultanément, en interaction, selon des boucles de rétroaction constantes et parfois contradictoires, dans la cadre de sessions de navigation parallèles qu'il faudra ensuite remettre en cohérence. De plus, chaque outil (moteurs de recherche, agents intelligents, annuaires ...) dispose de niveaux de structuration en général assez faibles et de toute façon non-homogènes<sup>265</sup> empêchant qu'on leur applique à chaque fois et à coup sûr les bons critères<sup>266</sup>.

Nous choisissons donc d'aborder les questions soulevées par les pratiques de browsing et celles de searching non plus sous l'angle exclusif des activités de recherche d'information mais de donner à ces concepts une résonance élargie pouvant rendre compte de manière opératoire de l'ensemble des stratégies de navigation possibles dans un environnement hypertextuel – invariants – confirmant les hypothèses et les conclusions déjà établies dans ce travail.

L'intérêt de ces deux notions venant directement de leurs points de similarité et de divergence, nous les présentons simultanément, en reprenant la liste de critères servant à les caractériser, établie dans le tableau ci-dessous.

<sup>264</sup> d'autant que, même pour les processus classiques dans lesquels le besoin en information est supposé constant : « *l'observation des conduites réelles des usagers [a montré que] le besoin en information ne reste pas constant au cours de l'interrogation.* » [Kolmayer 98 p.90]. Voir aussi dans le chapitre trois, le point 7.4.2. « Scénarios d'usage. »

<sup>265</sup> voir le point 7.4.

<sup>266</sup> d'autant que la plupart des outils partagent pour des raisons techniques et commerciales tout ou partie de certaines bases de données (ou bases d'index) servant à alimenter les différents annuaires ou moteurs de manière totalement indécélable pour l'utilisateur lambda (sur ce sujet voir les sites <http://www.abondance.com> et <http://www.searchenginewatch.com>).

fondement ni utilité. L'avantage offert dès lors par les ontologies, est de permettre de répondre à toutes les demandes formulées par n'importe quel membre à l'aide du vocabulaire partagé. Il ne devient plus indispensable d'avoir recours à un expert et des niveaux d'expertise différents peuvent cependant être maintenus<sup>90</sup>. [Gruber 93] « *Pragmatiquement, une ontologie commune définit le vocabulaire avec lequel les demandes [posing queries = ask] et les assertions [making logical assertions = tell] vont être échangées entre les différents agents (humains et/ou logiciels et/ou hybrides)* ».

#### **7.4. Niveaux ontologiques présents dans FoRSIC.**

Dans le projet FoRSIC, l'univers de discours est celui des pratiques de formation à la recherche d'information.

Le domaine comprenant les concepts est celui de la recherche d'information<sup>91</sup>. C'est le premier point que nous développerons ici.

La « formation » à la recherche d'information semble par nature se prêter à une mise en forme ontologique de type « orientée-tâche ». Chaque pratique de formation repose en effet :

- sur plusieurs éléments assimilables à des objectifs ou à des intentions pédagogiques et mettant en œuvre des niveaux de formation différents (1),
- sur des activités cognitives mobilisées pour remplir ces tâches (2),
- et sur un certain nombre de fondamentaux qui rendent compte des spécificités des pratiques de formation dans ce champ donné (3). Nous avons baptisé ces fondamentaux « actes documentaires ».

Le deuxième point que nous présenterons détaillera la manière dont nous avons mis en concordance l'ensemble de ces trois éléments, autour de scénarios d'usage.

Nous entrerons ensuite dans la présentation de la « matrice de compétences » utilisée dans FoRSIC : il s'agit du niveau ontologique qui fait l'interface entre le modèle de domaine et les scénarios d'usage. Le projet étant centré sur une communauté de formateurs, le modèle de domaine du champ de la formation (recherche d'information) n'est que l'un des aspects du vocabulaire partagé qu'il faut pouvoir proposer comme entrée ontologique pertinente aux membres de ce collectif. Nous avons choisi de construire cette matrice ontologique interface autour de la notion de « compétence en formation » : elle peut être utilisée par des experts (« je sais faire *telle chose* ») ou par des non-experts (« je souhaite faire *telle chose* »).

Nous relaterons enfin la méthodologie et les raisons qui nous ont ensuite conduit à retravailler sur cette matrice ontologique interface, pour distinguer entre le référentiel de compétences génériques qu'elle permet de produire automatiquement (pratiques de formation « idéalisées ») – et qui est au final interfacé de manière totalement transparente pour l'utilisateur – et la reformulation de ce référentiel en compétences explicites (liées aux pratiques de formation ayant effectivement cours dans le champ), lesquelles

<sup>90</sup> FoRSIC rassemble un collectif de formateurs aux niveaux d'expertise différents, certains, par exemple, spécialisés dans l'interrogation de bases de données en ligne et d'autres maîtrisant parfaitement les techniques de catalogage. L'utilisation d'ontologies permet aux experts d'exprimer leur savoir-faire et aux non-experts de faire part de leurs besoins, à l'aide de la même chaîne de caractère, des mêmes concepts.

<sup>91</sup> domaine incluant celui de la recherche documentaire. Ce modèle de domaine est disponible en annexe 11.

concepts. Ces deux derniers critères sont des standards de l'approche mathématique fréquemment utilisée en informatique<sup>87</sup> ;

- critère du biais minimal d'encodage (« *minimal encoding bias* ») : une ontologie doit être indépendante du niveau symbolique des termes du domaine dont elle rend compte ;
- critère d'extension monotone maximale (« *Maximum Monotonic Extendibility* ») : c'est l'un de ceux qui posent le plus de difficultés. Il doit permettre d'anticiper l'utilisation du vocabulaire partagé. Or la diversité des pratiques et la dynamique des concepts pour un champ de savoir donné, sont difficilement prévisibles. Il s'agit donc – c'est l'optique que nous avons choisie dans FoRSIC – de s'en tenir au vocabulaire et aux pratiques dont l'usage est avéré, en laissant ouvertes certaines entrées pour pouvoir les compléter à l'arrivée de nouveaux termes ou concepts ;
- critère de distinction ontologique : à partir du moment où une ontologie se compose de plusieurs classes – ce qui est le cas le plus fréquent – celles-ci doivent être disjointes<sup>88</sup> ;
- critère de diversification des hiérarchies : il permet la mise en place des processus constants d'héritage entre concepts d'une même classe (« *multiple inheritance mechanism* ») ;
- critère de minimisation de la distance sémantique entre concepts semblables : les concepts similaires doivent être représentés en utilisant les mêmes primitives, c'est-à-dire les mêmes relations (si « A » est un outil et si « B » est similaire à « A » alors « B » est un outil).

L'attention portée au strict respect de ces critères constitue une garantie méthodologique suffisante et reconnue. Les ontologies ainsi construites disposent de fonctionnalités leur permettant de faciliter le partage d'information, d'offrir une compréhension et des représentations communes d'un domaine, de faire coopérer, sur la base de cette compréhension réciproque, des personnes, des systèmes informatiques et des organisations, et surtout, comme c'est le cas dans le projet FoRSIC, de rendre explicites un ensemble d'usages, de pratiques et de terminologies de référence, habituellement implicites.

Du point de vue des usagers, ces ontologies permettent de fédérer la communauté au sein de laquelle ils s'inscrivent : en mettant à disposition de tous un vocabulaire partagé par tous, elles ne présupposent ni n'obligent à ce que chaque individu formant la communauté dispose d'une connaissance de l'ensemble des règles et des savoirs que recouvre ce vocabulaire<sup>89</sup>. D'autant que chacun sait – heureusement – des choses que les autres ignorent, faute de quoi la notion même de collectif ou de communauté d'usage n'aurait ni

<sup>87</sup> [Hofstadter 85 p.114] « Si la consistance est la condition minimale pour que les symboles acquièrent des significations passives, la notion complémentaire, la complétude, est la confirmation maximale de ces significations passives. Alors que la consistance d'un système formel est la propriété selon laquelle « toute assertion engendrée par le système est vraie », la complétude est l'inverse : « le système engendre toute assertion vraie. » »

<sup>88</sup> ce critère fait référence à la théorie des types logiques de Russell. [Bateson 80 p.10] « La thèse centrale de [la théorie des types logiques] consiste à dire qu'il existe une discontinuité entre la classe et ses membres : la classe ne peut pas être membre d'elle-même, pas plus qu'un de ses membres ne peut être la classe, et ce parce que le terme utilisé pour la classe ne se situe pas au même niveau d'abstraction que celui qu'on utilise pour ses membres. Autrement dit, il appartient à un autre type logique. » Nous avons, dans FoRSIC prêté une attention particulière au respect de ce critère qui conditionne également les différents niveaux « d'emboîtement des échelles » s'appliquant à tout type d'organisation hypertextuelle.

<sup>89</sup> à titre d'exemple, les membres du collectif dans le projet FoRSIC connaissent tous l'existence des termes « moteur de recherche » et « annuaire de recherche » sans pour cela que chacun en saisisse les différences et soit capable de les développer dans le cadre d'une ressource de formation.

	BROWSING (cruising)	SEARCHING (surfing)
Séquentialité ?	Top-Down (descendant/ascendant)	Bottom-up (émergent)
	« From where to what » Navigation ordonnée (itération)	« From what to where » Navigation intuitive (itinération.)
Dimension fractale.	« Effet Joseph » (persistance du contexte)	« Effet Noé. » (désorientation)
	Néguentropie	Entropie
Dynamiques.	Logique de fission	Logique de fusion
Styles cognitifs.	« Serialists »	« Holists »
	« field-independant »	« field-dependent »
Rapport aux invariants de liaison	Liens extensifs	Liens intensifs
	Ancre structurelles	Ancre inférentielles

Tableau 11 : Vues comparées des invariants de navigation.

### 7.7.1. Séquentialité ?

Le browsing se réfère à une activité de navigation qui s'inscrit dans une dimension séquentielle, ou dont l'objectif est de préserver une séquentialité d'accès à l'information recherchée. Il est caractéristique, par exemple, de la consultation systématique de tous les liens présents dans une table des matières hypertextuelle, et tend à préserver les niveaux d'arborescence qui y sont présentés. [Lazarder et al. 00 p.577] « (...) Le browsing hypertextuel fait appel aux mêmes compétences que la lecture d'un texte imprimé et l'utilisation d'aides à la recherche comme les index. »

A l'inverse, le searching est une approche plus émergente, une navigation capable de faire sens<sup>267</sup> en dehors de toute arborescence pré-établie. Il procède d'une navigation plus intuitive qu'ordonnée. Il est une « itinération » ouverte, centripète, quand le browsing est une itération fermée, centrifuge. L'analogie avec le jeu d'échec proposée par [Hofstadter 85 p.320] est ici tout à fait éclairante :

« Toute personne ayant un peu joué aux échecs a organisé sa perception de telle sorte que les déplacements des tours en diagonale, les prises frontales par des pions, etc... ne lui viennent jamais à l'esprit. De même, les joueurs de plus haut niveau se construisent, eux, des niveaux d'organisation supérieurs dans leur perception de l'échiquier. Il en résulte qu'il est aussi improbable que les mauvais coups leur viennent à l'esprit que des coups interdits à l'esprit du commun des joueurs. Cela pourrait s'appeler l'élagage implicite de l'arbre de branchement géant des possibilités. Par opposition, l'élagage explicite désignerait la réflexion sur un coup et, après examen superficiel, la décision de ne pas poursuivre plus loin cet examen. »

### 7.7.2. Dimension fractale.

[Gleick 91 p.125] à propos de la théorie des fractales de Mandelbrot, rappelle l'une de ses axiomatiques fondamentales qui isole deux séries d'effets : effet Joseph et effet Noé. « L'effet Noé signifie la

<sup>267</sup> dans la double acception du terme : signification et orientation.



*discontinuité : lorsqu'une grandeur varie, sa variation peut être presque arbitrairement rapide. L'effet Joseph, lui, correspond à la persistance. »*

Au vu des rapports entretenus par les activités de browsing et de searching avec la notion de séquentialité et d'arborescence, on peut considérer que l'effet Joseph s'applique de manière parfaitement adéquate au browsing, pour lequel l'important est, du point de vue de l'utilisateur, d'assurer et de maintenir une cohérence et une persistance du contexte cognitif ayant présidé à la navigation (néguentropie). Le searching relève lui de l'effet Noé, qui s'il comporte plus de risques de désorientation (entropie) favorise et accroît en parallèle les effets de sérendipité associative.

### 7.7.3. Dynamiques.

[Hofstadter 85 p.746] définit ainsi les dynamiques relevant de la fission et de la fusion :

*« La fission et la fusion sont deux idées intéressantes et complémentaires sur l'interaction des symboles. La fission est le détachement graduel d'un nouveau symbole [lien] de son symbole [lien] mère (c'est-à-dire du symbole ayant servi de modèle pour le produire). La fusion, c'est le phénomène qui se produit quand deux symboles [liens] ou plus, à l'origine sans rapport, participent à une « activation conjointe » en se transmettant si activement des messages l'un à l'autre qu'ils deviennent liés et que leur combinaison peut ensuite être adressée comme s'il n'y avait qu'un seul symbole [lien] ».*

Il suffit ici de remplacer le terme de « symbole » par celui de « lien » pour pouvoir inférer, en toute cohérence, que le browsing relève d'une dynamique de « fission » et le searching, d'une dynamique de « fusion »<sup>268</sup>.

### 7.7.4. Styles cognitifs.

La distinction opérée entre browsing et searching – confirmant en cela sa nature d'invariant – permet également de rendre compte de celle proposée par Horn et reprise par [Carter 97 p.26] distinguant deux catégories de lecteurs (holistes et sérialistes) en fonction des styles cognitifs<sup>269</sup> qui les caractérisent.

*« Horn approche le problème de la lecture du point de vue de la manière dont l'état d'esprit d'un lecteur s'accorde à un hypertexte. Il y a, écrit-il, deux genres de lecteurs : les sérialistes et les holistes.*

*« Les sérialistes éprouveront des difficultés avec l'hypertexte. Ils seront froissés d'avoir à faire des choix forcés. En voyageant au travers de nouveaux liens ils seront sujet à la confusion et à la désorientation beaucoup plus facilement que les holistes.(...) Ils s'offusqueront de l'introduction de nouveaux mots et de nouveaux concepts qu'ils ne comprennent pas. Ils n'accepteront pas de grandes « cartes conceptuelles » sur un sujet donné parce qu'elles contiennent des termes qu'ils ne connaissent pas. Ils bénéficieront des liens de définition plus que les holistes, parce qu'ils leur permettront de satisfaire leur besoin de certitude au sujet du sens des mots qui leurs sont peu*

<sup>268</sup> en consultant la suite de l'argumentaire d'Hofstadter, on pourra noter l'écho que ces notions entretiennent avec les problématiques afférentes au versioning et à la transclusion : « La fission est un processus plus ou moins inévitable, puisqu'une fois qu'un nouveau symbole a été réalisé par copie d'un ancien, il peut voler de ses propres ailes et ses interactions avec le monde extérieur se reflètent dans sa structure interne privée ; ce qui était, au départ, une copie parfaite, devient donc bientôt imparfait et de plus en plus différent du symbole ayant servi de modèle. »

<sup>269</sup> nous reprenons la définition de Riding & Rayner, 1998 citée par [Chen & Macredie 02] à propos des styles cognitifs : « l'approche habituellement préférée par un individu pour organiser et représenter l'information. »

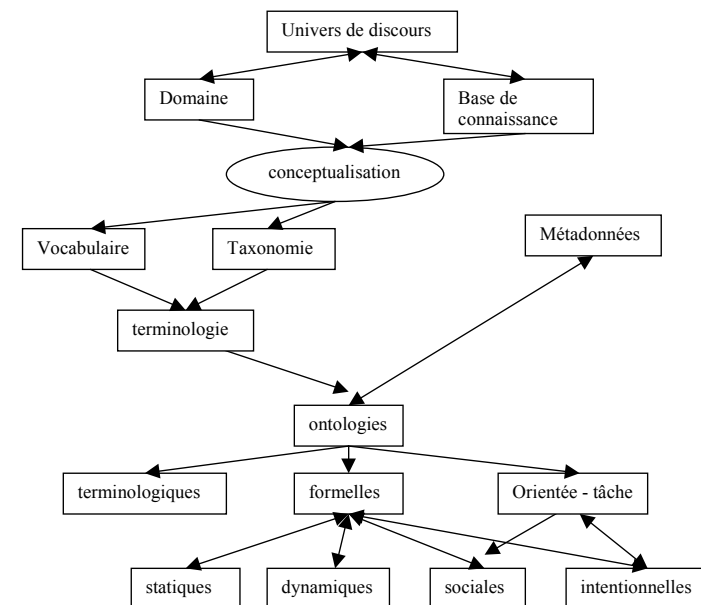


Fig. 22 : Types ontologiques.

### 7.3. Méthodologies d'élaboration et résultats attendus.

Avant d'entrer dans le détail de l'utilisation de tel ou tel type d'ontologie dans le cadre du projet FoRSIC, nous présentons les critères génériques définis par [Gruber 93] qui s'appliquent à la conception de n'importe quelle catégorie d'ontologie et en profitons pour signaler quelques-unes des principales fonctionnalités que la validation de ces critères dans le cadre de FoRSIC nous a permis de mettre en place.

Ces critères sont les suivants :

- critère de clarté et d'objectivité : il doit permettre de pouvoir communiquer à tous le sens voulu des termes définis dans l'ontologie;
- critère de consistance : il vise à ne proposer que des règles d'inférence qui soient cohérentes avec l'orientation choisie (c'est-à-dire avec le sens des termes disponibles). Une ontologie disposant des termes « quadrature » et « cercle » ne doit pas autoriser leur activation conjointe ;
- critère de complétude : il permet qu'une définition soit exprimée par des conditions nécessaires et suffisantes, afin d'éviter, notamment, les problèmes liés à la synonymie ou à l'homonymie entre

- les ontologies dynamiques s'intéressent aux aspects changeants d'un monde donné. Elles disposent la plupart du temps de trois concepts primitifs : « états », « états transitionnels » et « processus »<sup>86</sup>. Elles accordent une large place au repérage de relations temporelles (avant, après, pendant ...) déterminantes pour comprendre la nature des flux à l'œuvre dans tout système prétendant fonctionner de manière synchrone. [Jurisica et al. 99 p.486] font par ailleurs remarquer que la causalité (utilisée comme relation fondamentale dans nombre d'ontologies) est, dans ce cadre, une notion très proche de celle de temporalité : « elle impose des contraintes d'existence à des événements (si A cause B et que A vient de se produire, on peut s'attendre à B) » ;
  - les ontologies intentionnelles se rapprochent des ontologies de tâche. Elles englobent l'ensemble des « motivations, des intentions, des buts, des alternatives, des choix, des croyances et opinions diverses des agents de ce monde (je crois, je veux, j'approuve ou désapprouve ...), organisées en tâches et sous-tâches ». A ce titre, ce type a fait l'objet d'un grand nombre de travaux de modélisation dont on trouve la référence dans l'article de [Jurisica et al. 99 p.487] ;
  - les ontologies sociales, enfin, s'intéressent à la conceptualisation et à la modélisation de tous les aspects de type sociaux (« positions, rôles, autorité, permanence des structures d'organisation, réseaux d'alliances ... » [Jurisica et al. 99 p.489]). Cette dernière catégorie est l'une des plus complexes à mettre en œuvre, du fait du niveau d'implicite qui affecte chacun de ses aspects. Elle témoigne de l'importance qu'il faut accorder à la notion déjà évoquée de LPP et indique qu'il est « vital de cerner le contexte social pour identifier et sélectionner les réponses, les solutions techniques appropriées. » [Jurisica et al. 99 p.489] Bien que très complexes, ces ontologies sociales constituent souvent « l'unique moyen d'identifier les situations d'utilisation pertinentes et du même coup les réponses ou les méthodes à « favoriser ». » [Jurisica et al. 99 p.489]
- Au vu de cette typologie, nous proposons une extension à la figure 21.

<sup>86</sup> [Jurisica et al. 99 p.486] précisent que « Depuis les années 60, on a souvent utilisé des réseaux de Petri comme outils de modélisation appropriés pour des processus dynamiques discrets, notamment dans de nombreuses applications en télécommunications (applications temps-réel). »

*familiers .. Les holistes adoreront l'hypertexte. Ses possibilités de navigation s'accordent parfaitement à l'inclinaison naturelle des holistes. » »*

Le browsing favorise la mise en place de styles cognitifs « sérialistes » le searching relevant d'une approche « holistique »<sup>270</sup>. L'article de [Chen & Macredie 02 p.8] détaille la correspondance existant entre ces deux approches et les notions de navigation « field-dependent » et « field-independent » : « Les apprenants « indépendants d'un champ » font un meilleur usage des index pour localiser un item. (...) Les apprenants « dépendants du champ » préfèrent utiliser des cartes pour disposer d'une vue d'ensemble du contexte. », rappelant l'importance de prendre en compte ces paramètres pour proposer et implémenter des systèmes de liaison « adaptatifs »

#### **7.7.5. Rapport aux invariants de liaison.**

Notons enfin que ce qui nous conduit à poser le browsing et le searching comme invariants de navigation est le rapport qu'ils entretiennent avec les invariants de liaison présentés plus haut (propriétés des ancrés). Le browsing relève des propriétés contextuelles associées à des modes de création structurels quand le searching concerne les aspects inférentiels de ces mêmes modes.

De même le browsing est la traduction dans la navigation de propriétés collectives relevant de structures arborescentes (ancres extensives) quand le searching utilise l'éventail des mêmes propriétés collectives des ancrés, cette fois liées à des structures non-arborescentes (ancres intensives).

#### **7.7.6. Vers une troisième voie : navigations émergentes.**

Au vu de quelques-unes des pistes de recherche actuellement les plus explorées – interfaces haptiques, réalité virtuelle ou augmentée, holographie – commence à se faire jour dans la littérature un besoin de caractériser différemment les stratégies de navigation pouvant en rendre compte. La distinction entre browsing et searching reste communément admise, exception faite de quelques variantes terminologiques<sup>271</sup>, mais à chaque fois qu'il s'agit de décrire une navigation s'inscrivant dans l'une de ces voies, la question des limites du concept même de navigation se trouvent posées. Les stratégies d'orientation mises en place de manière immersive lorsque l'on entre dans la « peau » virtuelle d'un avatar et que l'on progresse dans un paysage de données (datascape) relèvent-elles encore d'une navigation ou procèdent-elles d'une nouvelle présence au monde (virtuel) ? Le terme de « questing » (être en quête) et sa référence explicite à l'univers des jeux de rôles qui ont constitué le fer de lance de ces applications pourrait alors se

<sup>270</sup> sur ces questions nous nous référons également à l'article de [Ford 00].

<sup>271</sup> [Burbules 97] propose ainsi de distinguer entre « surfing » ( « surfing » est un terme qui fait allusion à un groupe particulier d'utilisateurs, caractérisés par leur sexe, leur classe sociale, leur région et leur culture ») et « cruising » (« pour le même phénomène (qui bien sûr se réfère à des groupes différents) ») L'analogie searching/surfing résiste à l'analyse, et l'idée de « cruising » connotant une navigation de type croisière en laquelle tout serait établi à l'avance et où l'utilisateur n'aurait pas les commandes, complète également bien l'idée de browsing. Cependant les critères sociologiques choisis pour articuler cette distinction ne nous satisfont pas.

révéler plus approprié, cette quête pouvant prendre selon les cas la dimension d'une errance (« wandering »), ou à tout le moins d'une « dérive »<sup>272</sup>.

<sup>272</sup> [Abendroth & Decock 00b] « La comparaison de ces deux types de mouvement [browsing/searching] révèle la compréhension « psychogéographique » de l'espace développée par les situationnistes, opposant à l'emploi utilitaire de l'espace la stratégie subversive de « dérive ». « dérive : un mode de comportement expérimental lié à la condition urbaine : une technique de passage hâtif à travers des ambiances variées. Utilisé également, en particulier, pour désigner la durée prolongée d'une telle expérience. » <http://www.nothingness.org/SI/journaleng/IS1/definitions.html> Le concept de dérive est indissolublement lié à la reconnaissance d'effets de nature psychogéographique, et à l'affirmation d'un comportement ludique constructif, ce qui l'oppose en tous points aux notions classiques de voyage et de promenade. « ... La part d'aléatoire est ici moins déterminante qu'on ne le croit – du point de vue de la dérive, il existe un relief psychogéographique des villes, avec des courants constants, des points fixes et des tourbillons qui rendent l'accès ou la sortie de certaines zones fort mal aisé. » Théorie de la dérive. Guy Debord. 1956. »

En termes d'avantages offerts par les ontologies elles « *fournissent une compréhension commune et partagée d'un domaine qui peut être communiquée entre les gens et entre diverses applications* » et ce à partir du moment où « *elles incluent toutes un vocabulaire de termes et une spécification du sens de ces termes.* » [Carr et al. 01]

Par contre, elles ne sont valables en l'état que pour un domaine de connaissance donné et qui peut être circonscrit par une conceptualisation<sup>83</sup>. La nature des relations et des règles d'inférence alors utilisées et mises en place va permettre de déterminer des types ontologiques différents.

Le premier de ces niveaux distingue entre ontologies « terminologiques » et ontologies « formelles ». Souvent utilisées en complémentarité, elles n'en constituent pas moins des catégories différentes : « *la différence entre une ontologie terminologique et une ontologie formelle est une différence de degré : plus on ajoute d'axiomes à une ontologie terminologique, plus elle a de chances d'évoluer vers une ontologie (...) formelle.* » [Sowa 01]

Le second niveau concerne une catégorie particulière d'ontologies dans lesquelles la finalité consitue le premier critère de constitution : il s'agit des ontologies orientées-tâche (task-oriented). On en trouve un descriptif dans [Mizoguchi et al. 96]. Ce type d'ontologie est particulièrement utilisé en raison de sa portabilité (similarités possibles entre plusieurs domaines de connaissance<sup>84</sup>).

« Une ontologie de tâche est un système/une théorisation d'un vocabulaire pour décrire la structure inhérente à la résolution de problèmes de toutes les tâches existantes d'un domaine, de manière indépendante. La conception d'une ontologie de tâche a pour but de venir à bout des imperfections des tâches génériques tout en préservant leurs philosophies de base. Le but ultime de la recherche concernant les ontologies de tâche est d'offrir un cadre théorique pour l'ensemble du vocabulaire nécessaire à la construction d'un modèle du processus humain de résolution de problèmes

- Une ontologie de tâche s'organise autour de quatre genres de concepts :
- des noms génériques représentant les objets et reflétant leurs rôles tels qu'ils apparaissent dans le processus de résolution de problèmes.
  - Des verbes génériques représentant des unités d'action apparaissant dans le processus de résolution de problèmes.
  - Des adjectif génériques permettant de modifier les objets.
  - D'autres concepts spécifiques à la tâche. »

On trouve enfin la distinction opérée par [Jurisica et al. 99]<sup>85</sup> entre ontologies « sociales », « statiques », « dynamiques » et « intentionnelles ». Il s'agit cette fois d'une distinction de nature autant que de fonction :

- les ontologies statiques s'intéressent aux choses existant dans le monde – dans un monde donné – pour déterminer quels peuvent être leurs attributs et leurs relations. Il s'agit d'un type basique, qui ne prétend être ni universel ni trivial ;

<sup>83</sup> le web sémantique ayant vocation à systématiser l'utilisation des ontologies à l'ensemble des ressources présentes sur Internet.

<sup>84</sup> [Rastier 01] « Et si la connaissance était une action oubliée ? De même qu'un encyclopédie est une archive de passages de textes décontextualisés, une ontologie pourrait être définie comme une archive d'actions (...). » Cette proposition de Rastier illustre parfaitement l'approche ontologique retenue dans FoRSIC (point 7.4.) et fait le lien avec la notion d'archive foucauldienne (point 2.3.).

<sup>85</sup> Cet article est un état de l'art des principaux concepts utilisés en informatique pour représenter le savoir. Il analyse leurs avantages et inconvénients respectifs et permet de relier ces approches à la science de l'information (d'un point de vue théorique et pratique).

qui peut exister entre le thesaurus et la taxonomie. Nous choisissons de nous en tenir à la perspective évoquée plus haut selon laquelle un thesaurus se compose d'un vocabulaire auquel est associé un ensemble de règles permettant d'établir une taxonomie.

Notons enfin que les ontologies et l'ensemble des termes évoqués ici relèvent d'un univers de discours, c'est-à-dire d'un ensemble d'objets (individus, ressources, savoirs ...) qui constituent le savoir d'un domaine et qui peuvent être représentés dans un formalisme de type déclaratif [Gruber 93]. Cet univers de discours dispose à son tour de deux entités liées : un domaine dont va tenter de rendre compte l'activité de conceptualisation<sup>82</sup>, et une base de connaissance, qui comprend l'axiomatique permettant de représenter les relations existant entre les termes du domaine (il peut s'agir de règles sémantiques, logiques ou combinant les deux aspects).

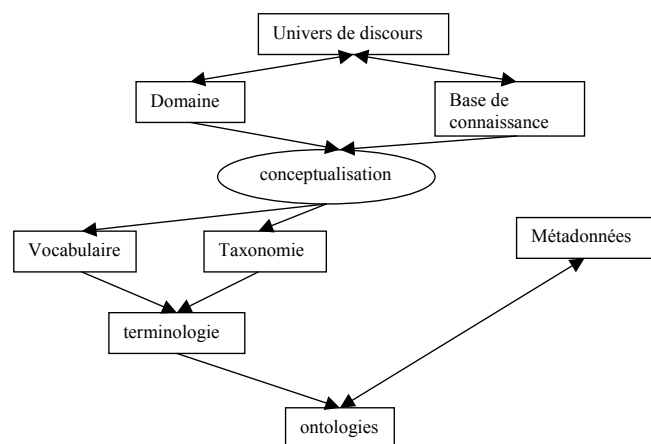


Fig. 21 : Eléments pouvant composer une ontologie.

### 7.2.2. Types ontologiques.

La plupart des ontologies actuellement utilisées se réfèrent à l'utilisation conjointe de tout ou partie des composants présentés dans la figure ci-dessus. Ainsi [Carr et al. 01] indiquent que « La plupart des ontologies ont trois composants majeurs qui peuvent être utilisés selon des règles d'inférence : une taxonomie, des relations entre concepts et des axiomes (règles) ». C'est parce qu'elles reposent sur cette architecture à plusieurs niveaux, qu'elles peuvent prétendre répondre aux lacunes individuelles des métadonnées (tout le monde ne parle pas le même langage, n'a pas la même approche) et des terminologies (le même concept ne signifie pas la même chose pour tout le monde). Cette complémentarité leur accorde certains avantages en même temps qu'elle pose certaines limites.

<sup>82</sup> [Gruber 93] définit une conceptualisation comme un résumé, une vue simplifiée du monde que l'on souhaite représenter dans un but précis.

## 8. Du lien au lieu : penser le réseau ...

Les questions que soulèvent la nature réticulée de tous les modes d'organisation qui nous structurent et dans lesquels nous évoluons, qu'il s'agisse d'organisation sociale, politique ou littéraire sont innombrables dans leurs implications mais puisent tous leur force problématique dans la manière dont nous serons amenés à penser le réseau, c'est-à-dire à évaluer, dans le cadre de notre évolution, la nature d'une complexité d'un nouveau genre. Quand Pascal affirmait déjà « l'homme est un roseau pensant » en insistant sur le lien si particulier qui fait de la souplesse du roseau la source et l'origine de sa robustesse, un simple glissement paronomastique nous permet d'affirmer au même titre que l'homme est, par nature, un réseau pensant. Chaque avancée dans des champs de savoir aussi distincts que la biologie, la linguistique, ou la sociologie nous le confirme.

La diversité des questions n'est cependant pas nouvelle comme le prouve la simple déclinaison de l'assertion pascalienne. Les outils, les formalismes et les modèles définis dans ce travail, doivent à leur tour se mettre au service d'une meilleure compréhension des mécanismes en jeu dans l'évolution combinée des outils, des langages et des organisations sociales. Les aspects politiques que soulèvent chaque jour un peu plus la connaissance chaque fois plus accrue et chaque fois plus complexe de ces nouveaux modes d'organisation est de fait, profondément polémique, et véhicule – au sens propre – des enjeux à la hauteur de l'étendue qu'elle recouvre.

Le réseau, tel que nous avons tenté de le définir et tel qu'il se donne à lire à ce jour, est un ensemble complexe de pôles affirmés, implicites ou en émergence, dont les intersections et les embranchements résultent de processus complexes d'attraction et de répulsion. Ces deux principales forces et leurs déclinaisons (concentration et démultiplication, aspects centripètes et centrifuges, etc.) posent la question d'un « devenir » qui reste à définir. Deux optiques s'opposent. Celle de Proudhon tout d'abord pour qui :

« Sur le réseau en échiquier, réseau fédératif et égalitaire, des routes de terre et des voies navigables, est venu se superposer le réseau monarchique et centralisateur des chemins de fer, tendant à subalterniser les départements à la capitale, à faire d'une grande nation, jusqu'alors libre, un peuple de fonctionnaires et de serfs, et à donner le démenti aux lois les plus certaines de la science économique en général et en particulier de l'industrie voiturrière. » Proudhon. Cité par [Mattelart 97 p.167]

A l'exact opposé de cette conception fait écho celle de [Lévy 88 p.20] pour qui « La multiplication contemporaine des espaces fait de nous des nomades d'un nouveau style : au lieu de suivre des lignes d'erre et de migration au sein d'une étendue donnée, nous sautons d'un réseau à l'autre, d'un système de proximité au suivant. »

Qui peut dire aujourd'hui l'orientation que prendra le réseau numérique ? Comment statuer, avec quels critères, et à quelle échelle sur la prédominance de ses aspects fédérateurs ou centralisateurs ? Est-il encore pertinent d'opposer des logiques de « nomadisme » et de « servitude » alors que l'usage du lien

(social, métaphorique, cognitif, littéraire ...) autorise à chaque instant la multiplication des alternatives (aussi aléatoires fussent-elles) en même temps qu'il rend virtuellement impossible l'existence d'impasses au sein de l'espace qu'il tisse ? Le gigantisme de cet espace n'est pas tant le sien propre que celui des parcours qu'il engendre et autorise. La nature singulière de chacun d'eux semble naturellement plaider en faveur du nomadisme de Lévy en même temps que leurs contraintes technologiques et leur répartition à l'échelle planétaire sont des révélateurs forts de la servitude avancée par Proudhon. Faut-il craindre, comme semble le penser Tim Berners Lee<sup>273</sup> et à sa suite [Marchal 01] l'apparition d'un « effet de chréode »<sup>274</sup> résultant d'un usage apauvri ou standardisé des processus de liaison ? S'expose-t-on à voir ressurgir « l'effet Saint-Matthieu »<sup>275</sup> de la scientométrie ? Est-il d'ailleurs simplement possible d'envisager le jour où « *quand les fonctionnalités hypertextuelles feront partie intégrante de notre environnement informatique, les travailleurs de la connaissance accepteront et intégreront l'hypertexte dans leurs pratiques quotidiennes de travail.* » [Balasubramanian 94] ?

La question du Lien pose naturellement celle du Lieu, c'est-à-dire celle d'un espace qui est son propre contexte, qui ne peut être pensé qu'en contexte, dans lequel les significations qui s'y croisent n'ont comme horizon de déploiement que celui contingenté par l'espace qui autorise ces croisements.

L'organisation hypertextuelle, les rapports qu'elle permet de tisser entre mémoires, réseau et intelligence collective, fournit un angle d'approche unique pour une compréhension globale des mécanismes d'accès et de traitement de la connaissance « *dans de vastes, distribués et hétérogènes espaces informationnels, que nous avons baptisés des écologies dynamiques d'information* » [Pitkow & Recker 94]. Notre dernier chapitre devra permettre de montrer si ces « *écologies dynamiques d'information* » peuvent ou non s'inscrire dans un cadre plus englobant relevant d'une écologie cognitive. Il démontrera qu'elles permettent à tout le moins de poser les bases d'une pragmatique de la connaissance.

<sup>273</sup> « (...) il existe un risque plus spécifique au web, celui que certains internautes tombent dans une espèce de trou culturel, où ils s'enfermeraient en suivant uniquement les liens hypertextes des sites qu'ils ont déjà vu. Le web favorise les phénomènes de renforcement au profit de culture extrêmes, dont il devient difficile de s'échapper. » [Berners-Lee 00 p.64]

<sup>274</sup> « empruntant ce terme au biologiste Waddington, [Christian Allègre] notait avec justesse que, de même qu'un phénomène naturel tend à repasser par le même itinéraire (c'est la situation classique de l'ornière), sur la toile plus un site est vu, plus il tendra à l'être par un jeu des renvois démultipliés au même – soit un mécanisme préoccupant par lequel, à mesure que les pages augmentent, le voyage moyen (pensons au tourisme) s'uniformise et se standardise. Devant le danger d'une telle concentration de la collecte, on aimerait donc aussi plaider pour une stochastique de la navigation et de la conservation des sites (...). » [Marchal 01]

<sup>275</sup> « Certains documents sont sur-cités en fonction de l'effet Saint Matthieu qui veut qu'on prête plutôt aux riches. » [Courtial 90 p.49]

## **7.2. Ontologies versus terminologies.**

Afin de comprendre la manière dont peuvent être articulés ces différents niveaux d'approche (présents chez [Gruber 93] et plus globalement dans tous les mécanismes de conception d'ontologies), nous devons revenir sur des distinctions terminologiques reconnues et retenues par l'ensemble de la littérature en ce domaine. Deux ensembles se dessinent : les notions complémentaires à la mise en place d'ontologies d'une part, et différents « types » ontologiques d'autre part.

### **7.2.1. Terminologies, taxonomies, métadonnées, thesaurus.**

La plupart des systèmes qui mettent en œuvre des ontologies disposent de tout ou partie de ces « outils » complémentaires, qui peuvent être utilisés en amont ou en aval de l'ontologie choisie.

Les métadonnées tout d'abord, permettent de décrire le contenu des nœuds d'information en utilisant une syntaxe et un modèle commun et portable à tout type de document (nom de l'auteur, date de création, etc ...). Pour autant, « (...) *le marquage ne sert à rien tant que personne ne parle le même langage.* » [Carr et al. 01].

Les terminologies permettent d'établir, pour un domaine ou un champ de connaissance donné, un vocabulaire décrivant l'ensemble de son spectre, de ses notions. Parce que leur premier critère de validité est leur nature exhaustive, elles sont la plupart du temps établies par un collège d'experts et posent de fait des problèmes d'appropriation et d'accès aux utilisateurs non-experts. Là encore, « (...) *utiliser un vocabulaire commun ne sert à rien tant qu'il ne signifie la même chose pour personne.* » [Carr et al. 01]. Définies par [Mizoguchi et al. 96] comme « *un ensemble théorique d'étiquettes de concepts* », l'établissement de ces concepts ne devrait pouvoir se faire qu'après être parvenu à un consensus sur la signification et la portée de chacun.

Ces terminologies reposent la plupart du temps sur deux entités là encore différentes et utilisées en complémentarité : des listes de vocabulaire et des taxonomies. On reconnaît là la figure du thesaurus, une liste de termes hiérarchisés et disposant de relations sémantiques.

Un vocabulaire est tout simplement un ensemble, une liste de mots. Nombre de domaines disposent de leur propre « vocabulaire métier ». Ces vocabulaires sont souvent assimilés ou confondus avec les ontologies : chaque mot renvoie souvent à un concept opératoire dans le champ d'établissement du vocabulaire. Pourtant deux éléments manquent à ces vocabulaires pour en faire des ontologies : « *Le vocabulaire est dépendant du langage, il lui manque l'universalité. Il est faible dans la description des relations entre les termes qui le composent.* » [Mizoguchi et al. 96]<sup>81</sup>.

Les taxonomies permettent de pallier le manque de « *relations* » des vocabulaires. « *Une taxonomie est souvent une hiérarchie de concepts dans laquelle chaque lien est de type « est un » ou « partie de ».* » [Mizoguchi et al. 96] Un nouveau niveau de confusion est ici présent, témoignant du glissement sémantique

<sup>81</sup> nous voulons d'emblée nuancer la notion d'universalité. Nous verrons que les ontologies, telles que définies ici, ne sauraient prétendre à l'universalité qu'à certaines conditions.

est une spécification formelle, explicite d'une conceptualisation partagée. » Chacun des termes entrant dans le cadre de cette définition est ensuite précisé par [Gruber 93] :

- « *formelle* » signifie qu'elles sont interprétables pour un système informatique (« *machine-readable* »);
- « *spécification explicite* » : pour [Gruber 93], entrent dans le champ de cette spécification définie comme explicite, des concepts, des propriétés, des relations, des fonctions, des contraintes et des axiomes, tous « *explicitement définis* » ;
- « *partagée* », la notion de partage se rattache ici à celle d'un savoir consensuel (« *consensual knowledge* »);
- « *conceptualisation* » : ce dernier terme désigne un modèle abstrait « *d'un phénomène dans le monde*. »

Nous avons choisi de retenir cette définition parce qu'elle rend compte d'un ensemble de problématiques nécessairement mises en œuvre lors de l'élaboration de toute ontologie, éventail que ne permettent pas de distinguer d'autres approches comme celle de [Mizoguchi et al. 96] pour qui il s'agit d'un « *système de vocabulaire/concepts primitifs utilisés pour construire des systèmes artificiels*. »

La voie tracée par Gruber présente ainsi le double avantage de caractériser les distinctions – sur lesquelles nous reviendrons – entre ontologie, thesaurus, métadonnées, terminologies, etc. et d'offrir un cadre méthodologique problématique qui peut être résolu à l'aide de critères formels et logiques (voir partie suivante) eux aussi posés par [Gruber 93]. L'impérieuse nécessité d'établir ces critères est conditionnée par l'ambivalence de chacun des termes proposés dans la définition de [Gruber 93] :

- à quelle(s) condition(s), à partir de quel moment et sous quelle(s) forme(s) un système formel est-il interprétable par un ordinateur ? Doit-il reposer sur une modélisation de type logique ? L'ajout de métadonnées peut-il suffire ?
- en plus d'être définis de manière explicite, quelle est l'architecture permettant de disposer en un tout cohérent de concepts, de relations, de fonctions, etc ... ? Comment ces fonctions et autres relations doivent-elles être caractérisées et définies ? Comment établir et maintenir une adéquation entre elles et le monde ou l'objet dont elles doivent rendre compte ?
- si ce monde ou cet objet témoigne d'un savoir consensuel, à l'aide de quels critères poser ce caractère consensuel ? La notion même de savoir ou de connaissance consensuelle est-elle pérenne dans un environnement distribué de pratiques et d'usages comme l'est l'organisation hypertextuelle ?
- enfin, quel doit être le niveau, l'étendue, et quels doivent être les choix méthodologiques permettant d'aboutir à une conceptualisation cohérente pour un phénomène donné ?

L'exposé des niveaux et des types ontologiques présentés et retenus dans FoRSIC permet de dégager des éléments de réponse.

## Citations originales.

### Point 6. Typologie des hypertextes.

- [Ovanesbekov 96] « *which sort of information is contained in separate topics ? What is the « logic basis » ?* »
- [Ovanesbekov 96] « *Which type of a topic make the majority ? Which structure is reflected in the text and connects separate topics in one whole ?* »
- [Ovanesbekov 96] « *textbooks, scientific papers, explanatory dictionaries, userguide, encyclopedia.* »
- [Miles 95]
  - « *The logocentric text : a traditional linear text with a single and consecutive pathway established through its spaces.*
  - « *The arboreal text : a central linear text is provided with a branching structure of annotations, commentaries and discussions opening off this central thread (the tree).*
  - « *The rhizome text : a non-linear text that may or may not have various central nodes joined in multiple ways with other text spaces, these links are understood to be thematically determined and defined.* »
- [Drexler 95] « *A public system must be open to an indefinitely large community, scalable to large sizes, and distributed both geographically and organizationally.* »
- [Drexler 95] « *a system that shows users all local links (no matter how numerous or irrelevant) is bare hypertext. A system that enables users to automatically display some links and hide others (based on user-selected criteria) is filtered hypertext.* »
- [Miles 00] « *It is clear that literary hypertexts tend to emphasise the paradigmatic, while instrumental hypertexts emphasise the syntagmatic, and that axes are defined by the syntagmatic sequences provided and formed.* »

### Point 7. Stratégies de navigation

- [Bernstein 99] « *Fear of disorientation revived calls for hierarchy and prescribed reducing the role of links.* »
- [Bernstein 99] « *the end of the navigation problem.* »
- [Bernstein 99] « *instructors using hypertexts in college classrooms found that students rarely seemed to feel disoriented by the hypertext, and noted that disorientation, by disrupting preconceptions and demonstrating gaps in the reader's knowledge, has always been an important pedagogical device.* »
- [Hubrich 98] « *Like Mark Bernstein of Eastgate and many others, I believe that the whole « Lost in Hyperspace » issue is a non-issue : bad systems and bad writing disorient, but bad writing always has.* »
- [Drexler 95]
  - « *ease of tracing references (...).*
  - « *Ease of creating new references (...)*
  - « *Information structuring (...).*
  - « *Global views (...)*
  - « *Customized documents(...)*
  - « *Modularity of information (...)*
  - « *Task stacking (...)*
  - « *Collaboration (...).* »
- [Drexler 95] « *the evolution of knowledge in society.* »
- [Conklin 87 p.40] « *The additional effort and concentration necessary to maintain several tasks or trails at one time.* »
- [Davison 97 p.144] « *Another reason for user confusion is cognitive overload. A common instance for this is the « Embedded Digression Problem », when a user follows so many alternative chains of thought (encouraged by the links in the document) that the original goal is lost.* »
- [Kleinberg 98] « *the number of pages that could reasonably be returned as relevant is far too large for a human user to digest.* »
- [McManus et al. 96] « *a learner-controlled medium may not be effective for primary instruction. First, the learner does not have sufficient knowledge about the new content and therefore cannot make the best decisions about which concepts to study or what strategies to use. Second, the learner does not have the appropriate metacognitive ability to monitor his learning process. Third, the learner may not understand how to integrate his prior knowledge in the learning process.* »
- [Davison 97 p.144] « *There are two main reasons for a user getting « lost » while browsing a large web document. The first is a lack (or loss) of coherence ; this is also known as « The Art Museum Problem », which is the inability of the user to build any more abstractions about the document after browsing for a long period.* »
- [Young 90 p.240] « *provide a local view of the roles of individual links but give little indication of how the set to which the links belong is structured.* »
- [Young 90 p.239] « *helps users recognize and interpret what is contained in the node, and make decisions about what to do next more easily. (...) But internal node structure often gives little indication of how a collection of nodes are interrelated.* »
- [Young 90 p.240] « *Hierarchical skeleton, or tree structure : help users develop a mental model of the system and maintain orientation. (...) But a user has no way of knowing if links imply a part/whole relationship, or if nodes follow some logical breadth-first or depth-first sequence, or if information at the leaves of the tree support or elaborate on internal nodes.* »

- [Hofman & Oostendorp 98 p.123] « *The fact that a structural overview provides global structure was apparently outweighed by the fact that the structural overview leads the reader's attention away from the local structure of hypertext.* »
- [Carr et al. 99b] « *While tours often serve the purpose of introducing a subject, trails offer themselves to more advanced applications.* »
- [Balasubramanian 94] « *Hypertext provides the inherent capability of creating lateral classifications.* »
- [Lucarella 90 p.81] « *This paper approaches the problem of information retrieval from hypertext. In this context, the retrieval process is regarded as a process of inference that can be carried out either by the user exploring the hypertext network (browsing), or by the system, exploiting the hypertext network as a knowledge base (searching).* »
- [Lucarella 90 p.83] « *Browsing : going from where to what (presumably you know where you are in the database and you wanna know what is there). Searching : going from what to where (presumably you know what you want and wish to find where in the database it is).* »
- [Kleinberg 98 p.2] « *Our work originates in the problem of searching on the www, which we could define roughly as the process of discovering pages that are relevant to a given query.* »
- [Lazarder et al. 00 p.577] « *hypertext browsing calls on the same skills as reading printed text and using search aids as indexes.* »
- [Carter 97 p.26] « *Horn approaches the issue of readers from a perspective of how « matched » a reader's mindset is to reading hypertext. There are, he writes, two kinds of readers, the serialists and the holists :*  
     « *Serialists will have some difficulty with hypertext. They will resent the forced choices. Traveling to new links they will be thrown into confusion and disorientation much more easily than holists. (...) They will resent the introduction of words and concepts they don't understand. They will not accept large « conceptual maps » of the subject matter because they contain terms they don't understand. They will benefit from the definition links more than holists because they will be able to satisfy they need for certainty about the meaning of unfamiliar terms ... Holists will love hypertext. Its browsing capabilities are perfectly matched to the natural inclination of holists.* » »
- [Chen & Macredie 02] « *an individual preferred and habitual approach to organizing and representing information.* »
- [Chen & Macredie 02 p.8] « *Field-Independent learners made greater use of the index to locate particular item. (...) Field-Dependent learners prefer to use the map to get the whole picture of the context.* »
- [Burbules 97] « *« surfing » is a term that alludes to a particular group of users, particularized by gender, class, region, and culture » « cruising » for the same phenomenon (which, of course, alludes to a quite different group).* »
- [Balasubramanian 94] « *when hypertext functionality becomes an integral part of our computing environment will knowledge workers accept and incorporate hypertext into their daily work process.* »
- [Pikow et Recker 94] « *in large, distributed, heterogeneous information spaces, or what we have termed dynamic information ecologies* »

## 7. Le rôle à jouer des ontologies.

*« L'art d'arranger les choses du général au particulier n'est pas de moindre importance et nous est d'une grande aide pour notre capacité de jugement comme pour notre mémoire. (...) Cela ne nous aide pas seulement à retenir les choses mais également à les retrouver. » G.W. von Leibnitz, New Essays on Human Understanding. Cité par [Sowa 01].*

De l'arbre de Porphyre aux catégories aristotéliennes, des règles d'inférence qu'il développa et appliqua aux syllogismes, jusqu'à la logique « de premier ordre » [Sowa 01], des classifications du savoir de la période védique (1500-326 av JC) aux gigantesques thesauri actuellement disponibles (MESH lui-même inclus dans UMLS<sup>77</sup>) en passant les sept arts libéraux gréco-romains (organisés en Trivium et Quadrivium), les classifications du savoir proposés par Callimaque (240 av JC), Francis Bacon (1521-1626), Ranganathan, Dewey et tant d'autres<sup>78</sup>, les ontologies ont depuis toujours accompagné l'histoire de la connaissance, de ses modes de constitution et de ses possibilités d'accès. Elles disposent maintenant de leurs langages (KIF, RDF<sup>79</sup> ...), de leurs outils de génération (OntoWeb, OntoBroker ...), et de leurs champs d'application<sup>80</sup>.

Nous en avons déjà plusieurs fois fait mention dans ce travail, de manière allusive dans notre premier chapitre et plus opératoire dans le second, en présentant certains systèmes utilisant ces ontologies aussi bien comme principe de classification que comme technique de recherche d'information. A plusieurs reprises, nous avons laissé entendre, conformément à l'orientation que prône le web sémantique, qu'elles représentent l'une des voies actuelles de développement les plus prometteuses pour mettre en œuvre, dans le cadre de l'organisation hypertextuelle, une pragmatique de la connaissance. Nous voulons maintenant argumenter ce postulat en questionnant les deux faces de la réalité ontologique : leur rôle dans l'organisation de la mémoire et dans celle de la connaissance.

Nous présentons d'abord le cadre d'application de ces ontologies, depuis les premiers systèmes de classification les utilisant jusqu'aux derniers avatars du web sémantique.

Nous détaillons ensuite l'utilisation que nous en faisons dans le cadre du projet FoRSIC, les différents niveaux ontologiques qui le structurent, comment nous les avons mis au point et le type d'organisation (hypertextuelle) dont ils rendent compte.

### 7.1. Définition(s).

Les ontologies sur lesquelles nous allons baser notre argumentaire ne sont plus celles définies dans leur domaine d'origine, la philosophie, où elles désignent une théorie globale de l'existence. Nous empruntons la définition qui sera valable dans notre problématique à [Gruber 93] pour qui : « *Une ontologie*

<sup>77</sup> MESH : Medical Subject Headings, UMLS : Unified Medical Language System.

<sup>78</sup> Pour une perspective globale sur ces systèmes de classification et leur histoire on pourra consulter le remarquable diaporama consultable à l'adresse : <http://www.libsci.sc.edu/hoerman/clis732/Classi-2.pdf>.

<sup>79</sup> KIF : Knowledge Interchange Format, RDF : Ressource Description Framework. Sur ces langages et les outils mentionnés après (OntoWeb et Ontobroker), voir les nombreux renvois depuis le site du web sémantique (<http://www.Semantic-web.org>).

<sup>80</sup> innombrables, on citera parmi les principaux : la gestion et l'ingénierie des connaissances, l'intelligence artificielle, l'ingénierie linguistique – indexation automatique –, le web sémantique, la constitution des répertoires « à la yahoo » sur Internet, auxquelles s'ajoutent toutes les ontologies cachées dans divers systèmes de recherche et d'accès à l'information.



[Bateson 77 p.231] indique qu'« Une unité d'information peut se définir comme une différence qui produit une autre différence. Une telle différence qui se déplace et subit des modifications successives dans un circuit constitue une idée élémentaire. » Il pose du même coup toute différence circulante comme source de connaissance. C'est dans cette optique batesonienne qu'il faut lire la proposition de [Lévy 90 p.157] pour qui « Tout ce qui produira une différence dans un réseau sera considéré comme un acteur, et tout acteur se définira lui-même par la différence qu'il produit. »

Dans FoRSIC, tout élément du système (compétences, formateurs, ressources de formation, commentaires sur les compétences, etc.) peut être représenté en relation avec un ou plusieurs autres dans un arbre de connaissance. Les véritables acteurs du collectif ne sont donc pas simplement des individualités biologiques mais tout élément permettant de produire une différence repérable sur les cartographies produites, tout ce qui, occasionnant cette différence structurellement repérable, définit et devient une connaissance.

L'axiomatique qui pose que l'information est une différence et que cette différence existe à l'échelle du réseau en ce qu'elle occasionne une transformation dans la représentation du collectif impliqué, est une axiomatique fondamentale pour l'établissement d'une pragmatique de la connaissance. Comme annoncé plus haut<sup>76</sup>, FoRSIC permet d'atteindre l'adéquation « idéale » entre carte et territoire, faisant de l'organisation hypertextuelle un cadre opératoire pour la mise en œuvre d'une pragmatique de la connaissance. Les cartographies qui y sont proposées sont des cartographies de discours non pas indépendantes mais « au service », « rendant compte », de logiques territoriales perceptibles dans les dispositifs socio-organisationnels produisant ces discours.

# LE LIEU



« Je méditais sur ma demeure. Toutes les parties de celle-ci sont répétées plusieurs fois. Chaque endroit est un autre endroit. [Chaque endroit est] en nombre infini. La demeure a l'échelle du monde ou plutôt, elle est le monde. » [Borges 67 p.89].

« Que le langage ne détermine qu'en spatialisant (...). » [Derrida 67 p.28]

<sup>76</sup> voir le point 5.3. « Cartes fractales d'un territoire rhizomatique. » de ce chapitre.

## Section A.

□ **« Arbres de connaissances ».**

La philosophie des « arbres de connaissances » sur laquelle repose le logiciel SEE-K de la société Trivium fut élaborée par Michel Serres, Pierre Lévy et Michel Authier (actuel directeur de la société). On en trouvera une description exhaustive dans [Authier & Lévy 93]<sup>75</sup>.

*« Né comme concept sociologique et mathématique en novembre 1991, validé comme solution informatique en février 1992, proposé comme moteur d'un projet d'organisation des forces d'enseignement et de formation ouverte et à distance en avril 1992, exposé dans un livre en novembre 1992, devenu une réalité technologique en 1993 et industrielle en 1994, les " Arbres de connaissances " prétendent à un renouvellement des pratiques humaines aussi bien en situation de travail ou d'apprentissage, que dans la vie sociale ou éducative. Basés sur des principes nouveaux de traitement de l'information, de l'implication des acteurs et de l'exploitation des richesses humaines (...) » [Authier 98a]*

Principale dimension cartographique disponible dans le projet FoRSIC dont ils constituent le cœur de l'activité de représentation, nous revenons sur leurs propriétés, et expliquons pourquoi ils sont les techniques de représentation les plus adaptées pour rendre compte de tout type d'organisation hypertextuelle.

Si nous nous référons aux arguments et aux principes développés dans ce travail sur la problématique du versioning et celle de la transclusion, nous constatons une première adéquation de nature : la transclusion et le versioning constituent une tentative de réponse au problème de la compréhension et de la gestion des similarités non-identiques de contenus. La philosophie de la représentation véhiculée par les arbres de connaissance repose sur la même articulation : il s'agit d'offrir une lecture de la similarité et du rapprochement rattachés à une identité de représentation (forme des arbres) spontanément reconnaissable et toujours identifiable.

Les arbres de connaissance sont de fait la forme topologique la plus adaptée pour caractériser la nature de l'information comme « différence » dans un environnement hypertextuel.

Voyons d'abord ce qui peut être dégagé de leur caractérisation en tant que forme, au sens topologique du terme. Pour [Thom 93 p.35] « *Ce qu'on appelle habituellement une forme, c'est toujours, en dernière analyse, une discontinuité qualitative sur un certain fond continu.* » Cette discontinuité qualitative permet l'établissement d'une structure sous laquelle cette forme s'offre à la lecture et à l'interprétation. [Wittgenstein 61 p.33] « *La forme est la possibilité de la structure.* » Et ce qui, au sein de la structure ainsi produite, est source de signification, là où se loge le sens, ce sont les discontinuités, les brisures qu'elle permet d'identifier spontanément, si infimes et discrètes puissent-elles être. [Thom 93 p.63] « *Il en est souvent ainsi : le caractère discret d'une transformation est une simplification réalisée par notre appareil perceptif. Nous sommes faits pour voir essentiellement des discontinuités. Elles seules sont significatives.* » Les arbres de connaissance reposent sur ce principe : en évacuant – ou du moins en rassemblant – ce qui est massivement commun, ils permettent d'identifier, de repérer des différences et de comprendre leur nature, leurs significations.

<sup>75</sup> on pourra également consulter le site web associatif <http://www.globenet.org/arbor> sur lequel se trouvent des textes théoriques, des descriptions de projets les utilisant, etc.

Dans FoRSIC, chaque membre du collectif a la possibilité de décrire son organisme de rattachement (distance sociale), son environnement de formation, d'indiquer par exemple ses « personnes-ressources » (distance personnelle), de renseigner les parties concernant son état-civil (dimension publique) et d'exprimer un point de vue critique sur ses propres pratiques de formation, sur sa propre expérience (dimension intime).

### **6.5.3. Dimension cartographique et « Arbres de connaissances ».**

« (...) il fallait comprendre le savoir de chacun dans l'espace du savoir de tous. Ainsi chacun serait respecté et tous pourraient s'orienter. » [Authier 98b p.21]

Nouveau paradoxe – apparent – de FoRSIC : en s'efforçant de rendre compte de modes d'organisation hypertextuels, c'est bien la dimension collective qui sert d'interface, et non celle cartographique que nous allons maintenant détailler. Cette dernière permet de faire le lien, de confronter en un même espace de consultation, les logiques de coopération qui y sont opérantes et leurs modes de distribution, à l'échelle du collectif et des individualités le composant.

En « cartographiant » le collectif, on lui donne accès aux logiques de coopération qui configurent l'agencement des savoirs et des pratiques lui permettant d'exister comme entité homogène, et non comme simple addition de subjectivités. Ces logiques deviennent non seulement lisibles mais surtout interprétables et modifiables.

L'activité de représentation, notamment sous ses modes cartographiques, est au cœur même de l'organisation hypertextuelle – dont elle rend la topologie « navigable » – et au cœur des thématiques caractérisant les sciences de l'information et de la communication. Représenter c'est formaliser au moyen d'un certain nombre de règles, une réalité fuyante qu'il importe de rendre accessible dans son étendue et sa complexité. Dans FoRSIC, la réalité de la masse documentaire disponible (ressources de formation) n'est que l'un des éléments entrant dans le cadre de la cartographie.

« Une représentation consiste à faire apparaître à l'esprit un contenu. Les représentations sont produites et liées entre elles selon des modes réglés, dont je dirais qu'ils forment les modélisations primaires du savoir : le mode analytique, le mode symbolique et le mode pratique. (...) Ainsi trois modes fondamentaux : pratique, analytique, symbolique ; et trois fonctions constitutives : cognition, substitution, délégation. » [Reichler 89 p.98]

Nous allons maintenant rendre compte des techniques de représentation mobilisées dans FoRSIC et expliquer en quoi elles permettent d'intégrer ces trois « modélisations primaires du savoir », cette conjonction nous fondant à parler de cartographies de connaissances.

« L'espace virtuel déstabilise notre définition du réel et nos rapports perceptifs, cognitifs et symboliques avec l'environnement. » [Lefebvre & Pailliat 92 p.8]

Toutes les figures qui gravitent autour de l'hypertexte posent de manière récurrente la question centrale du lieu, chacune la déclinant en un nouvel espace, en une nouvelle topographie. Si après le livre et le lien, nous avons choisi de retenir le lieu comme troisième volet de notre trilogie hypertextuelle c'est parce que nous avons voulu procéder par emboîtements successifs, tout en ne privilégiant aucun sens particulier de lecture : le livre se donne à lire sous de nouvelles modalités au travers de l'usage des liens, qui mettent en place une topologie particulière de l'organisation de l'ensemble. Le lieu est structurellement perceptible et peut être investi par les parcours qu'autorisent les liens, renvoyant aux nouvelles formes du livre.

Les croisements de chacune de ces questions sont cependant différents : avec le livre, nous avons voulu sonder des transversalités de nature qui impliquent l'ensemble de la chaîne de production des discours (instances d'énonciation, genres, etc ...). Pour le lien il s'agissait principalement de transversalités liées à des fonctions données, qu'elles soient structurelles et/ou organisationnelles. Pour le lieu, il s'agit maintenant de faire état des croisements et des nouvelles transversalités qu'occasionne la présence, en un même espace, des deux précédentes.

Nous avons organisé ce dernier chapitre en trois parties.

Dans la première, nous abordons la question du lieu par deux de ses modalités caractéristiques de l'ère du numérique : l'écran pose la question du lieu d'apparition du texte à l'échelle individuelle et les bibliothèques sont les lieux d'organisation de nos savoirs et de nos connaissances collectives.

Sans écrans, pas de numérique<sup>1</sup>. Une fois franchi ce seuil qui est à la fois surface et interface, une fois entré dans ce monde, nous nous trouvons, que nous le voulions ou non, dans une organisation qui préexiste à notre entrée en ce monde (notre session de navigation) et qui survivra à notre sortie. Il s'agit donc d'un lieu, d'un espace, l'équivalent de nos bibliothèques, devenues, elles aussi, numériques. Après le passage du miroir-écran et la prise de repères dans ce nouveau « monde-bibliothèque » nous analyserons les nouvelles configurations qu'adoptent ces bibliothèques : après un rappel de leurs spécificités numériques, nous plaiderons en faveur de la mise en place d'une archivistique topologique pour rendre compte, de manière adéquate, de la connaissance accessible à travers elles.

Dans la seconde partie de ce travail, nous organiserons la problématique du lieu autour de deux axes : celui de la topologie comme horizon interdisciplinaire permettant de rendre compte de la plupart des typologies utilisées dans nos deux premiers chapitres, et verrons comment cette topologie particulière pose, à l'échelle de l'organisation hypertextuelle, la dialectique de la carte et du territoire.

Enfin, la dernière partie de ce chapitre sera consacrée au projet de recherche FoRSIC. Ce projet nous a permis de mettre en place un modèle pérenne de collaboration (social) et de construction de connaissances (cognitif) en utilisant les propriétés topologiques de l'organisation hypertextuelle pour mettre en œuvre une

<sup>1</sup> hors quelques approches à ce jour expérimentales (holographie notamment). Voir aussi annexe 4 « Stratégie des interfaces ».

pragmatique de la connaissance. Nous détaillerons notamment l'utilisation et la place qui est faite aux ontologies : celle de langages communs de classification, d'organisation, d'échange et d'émergence de nouvelles connaissances. Nous discuterons la thèse selon laquelle ces ontologies sont amenées à jouer un rôle moteur dans l'organisation du lieu de la connaissance à l'échelle du réseau, au travers d'un certain nombre de conditions nous semblant nécessaires et par la validation de croisements – encore « topologiques » – entre certains types ontologiques.

possible, ce qui implique qu'il puisse s'auto-organiser ou à tout le moins offrir à ses membres la possibilité d'identifier des situations non-coopératives<sup>73</sup>.

Comme l'énonce [Lévy 90 p.25], indépendamment du système ou de l'environnement socio-cognitif dans lequel elle prend place, « *La circulation d'information n'est souvent qu'un prétexte à la confirmation réciproque de l'état d'une relation.* » Ceci est vrai dans une perspective duelle, dialogique où les acteurs (émetteur et récepteur, auteur et lecteur, apprenant et enseignant ...) se trouvent en situation de co-présence. Dès qu'une tierce partie intervient – comme c'est le cas dans FoRSIC où toute situation d'énonciation, de circulation d'information prend place dans un cadre collectif – cet « *état de relation* » dont parle Lévy n'est pas seulement confirmé mais nécessairement reconfiguré : la dimension collective ne procède pas par addition des relations inter-individuelles, elle ne s'y substitue pas davantage, elle s'y sur-ajoute. Le tout est plus que la somme de ses parties.

### **6.5.2. Dimension collective.**

Si la dimension collective est à ce point structurante dans FoRSIC, c'est paradoxalement parce qu'elle est transparente pour l'usager : elle est présente comme arrière-plan commun à l'ensemble des situations de communication qui s'y déploient. Elle ne demande aucun effort ou aucune démarche volontariste. Elle est instituée de fait. Chaque élément nouveau de discours ou d'action, s'inscrit dans un territoire collectif (celui des pratiques de formation), concerne une communauté d'acteurs (individualités composant le collectif) et repose sur un « *quasi-objet* » qui recouvre et confirme l'horizon collectif ainsi établi : dans FoRSIC ce « *quasi-objet* » est constitué par le modèle de domaine de la recherche d'information. La réunion dans un cadre ensembliste de ces trois éléments<sup>74</sup>, définit la dimension collective de FoRSIC et plus globalement celle de toute organisation de type hypertextuelle :

*« Qu'est-ce qu'un collectif ? Une triade : communauté d'acteurs, territoire, quasi-objet. Sans quasi-objet, le territoire n'a pas de limite significative, les acteurs n'ont rien à faire circuler ; sans territoire la communauté est figée, le quasi-objet ne peut circuler ; sans acteur ...*

*Le quasi-objet ne peut être un morceau du territoire autrement c'est la guerre ; il ne peut être un acteur sans quoi le collectif devient victimaire. Il doit être non-humain comme le territoire et non-inerte comme les acteurs. Il doit bouger avec les acteurs sans avoir aucune aptitude autonome au mouvement comme le territoire. La cinématique des mouvements du quasi-objet détermine à la fois la taille du territoire et les dynamiques du collectif. » [Authier 98b p.225]*

Pour valider l'échelle des représentations collectives sur lesquelles chaque individualité peut choisir de se positionner, nous nous sommes efforcés dans FoRSIC de reprendre celle proposée par Hall comme une typologie des distances inter-personnelles :

*« Pour Hall, chaque culture organise l'espace de façon différente à partir d'un substrat animal identique, le « territoire ». Hall propose ainsi une échelle des distances interpersonnelles. Quatre distances sont envisagées : intime, personnelles, sociale et publique. Chacune comporte deux modalités : proche et lointaine. » [Winkin 81 p.89]*

<sup>73</sup> voir le point 7.6.4. « Invariants procéduraux et déclaratifs. » du chapitre deux.

<sup>74</sup> territoire collectif, communauté d'acteurs et quasi-objet sont pour partie inclus et pour partie éléments des deux autres.

*coopération est élevé, plus s'ouvre l'éventail des formes en compétition.* » [Lévy 00a pp.105-106]

La coopération avant d'être, ou non, un état de fait mesurable au vu des pratiques effectives ayant cours dans un dispositif, s'organise autour de principes théoriques structurants, permettant de la faciliter et de l'entretenir, c'est-à-dire de créer les conditions optimales de son amorçage, puis de son renforcement autour de pôles, d'acteurs ou de thématiques identifiés et repérables par tous.

Nous avons ainsi, dans FoRSIC, mis œuvre un axe paradigmatique de la coopération qui repose sur les principes énoncés plus haut.

Le principe de LPP constitue le premier repère de cet axe paradigmatique.

La cognition distribuée en est le second : en autorisant la propagation d'états de représentations à l'ensemble des acteurs présents dans le dispositif – notamment par le biais des représentations cartographiques – tout acteur ou ensemble d'acteurs constitué à un instant « t » représente et se représente le dispositif à l'intérieur duquel il est inclus.

L'analyse de la situation est le troisième de ces repères et repose sur trois principes distincts et complémentaires :

- principe de réflexivité : dans FoRSIC, la situation de partage, de mutualisation des connaissances et des savoirs peut en permanence être reconstruite *a posteriori*, à partir et sur la base des échanges qui l'instituent ;
- principe de descriptibilité : dans FoRSIC, on observe une concordance de chaque instant – c'est-à-dire demeurant au delà des limites temporelles de chaque session individuelle – entre action et discours. Chaque action (ajout d'une ressource, d'un besoin, etc ...) peut être assortie d'un discours en rendant compte, et chaque élément de discours présent dans le dispositif (texte de commentaire sur des compétences par exemple) rend compte d'une action de formation, de représentation ou de positionnement. Chaque nouveau discours renforce ou modifie la représentation et la place dans le collectif de l'acteur auquel il se rattache ;
- principe d'indexicalité : l'usage qui est fait des ontologies permet de rendre compte, d'indexer à l'aide de la même chaîne de caractères (référentiel de compétence) des textes, des situations et des individus.

Dernier de ces repères témoignant d'un processus coopératif, FoRSIC favorise l'auto-organisation en autorisant la gestion du processus émergent. En prenant en compte et en permettant de rendre compte en des termes équivalents mais non semblables, de facteurs sociaux, cognitifs, organisationnels et stratégiques, FoRSIC permet à la signification individuelle et collective de s'inscrire dans une circularité constructive laissant la place à l'émergence qui, une fois repérée, peut alors à son tour être intégrée dans la dynamique d'échange<sup>72</sup>. Ce dernier point est essentiel car pour être fonctionnellement adéquat, un système doit être capable de fournir des réponses correctes à son environnement dans un temps limité et aussi synchrone que

<sup>72</sup> Selon les principes du modèle de conversion de connaissances de [Takeuchi & Nonaka 95].

## 1. L'écran : le dialogue du corps et du texte.

« [L'écran a] le pixel pour alphabet, le signe pour vocabulaire, l'algorithmique pour grammaire et l'interactivité pour mode de lecture ». [Dall'Armellina 00]

L'écran, de manière récurrente et consubstantielle à sa nature, pose à différents niveaux la question du lieu. Comme nos deux premiers chapitres se sont efforcés de le montrer il est :

- un lieu de consultation : celui par lequel le lecteur/utilisateur entre dans l'hypertexte, celui qui scénarise et orchestre l'ensemble des modalités du regard qui s'y déploient (introspection, inspection, exo-spection) ;
- un lieu de manifestation : c'est au travers de l'écran que se donne à lire le comportement digital, cinétique du texte affiché ;
- un lieu « atomique » : le web est fondé sur la page-écran comme unité atomique de de navigation, laquelle prend souvent le pas, en termes de perception, sur la lexie ;
- un lieu de création : l'hypertexte, les dispositifs hypertextuels en général, sont construits et pensés pour un affichage écran ;
- un lieu dont la détermination topographique s'avère déterminante pour certaines des problématiques spécifiques de l'hypertexte : ainsi le problème de la transclusion ou celui du versioning sont directement liés à la taille des écrans.

Pour autant, la richesse et l'enjeu des problématiques qui le « traversent » tiennent à sa nature interface, c'est-à-dire aux rapports qu'il permet d'instituer entre deux subjectivités (auteur et lecteur) et entre le « discours » de l'une (hypertexte auctorial) et le parcours de l'autre (navigation lectorale). La logique du lieu qui le détermine est celle du seuil : face à l'écran, nous ne sommes pas encore vraiment dans l'interface mais nous ne sommes déjà plus uniquement dans le monde réel. « *Le jeu technologique avec la persistance rétinienne, base de la visualisation cathodique, nous fait basculer dans le monde du spectaculaire et du simulacre.* » [Laufer 89 p.220]. C'est dans ce monde que s'organise la dimension « panoramique »<sup>2</sup> de la littérature hypertextuelle.

Nous voulons ainsi relativiser le point de vue de [Vandendorpe 99 p.90] selon qui « (...) les contraintes de la lecture sur écran et la masse énorme d'informations accessibles ont amené la mise en place du concept d'hypertexte (...). Cette caractéristique situe l'hypertexte dans une pragmatique de l'interactivité. » Si l'écran relève bien d'une pragmatique de l'interactivité, l'hypertexte est, plus globalement, du côté d'une pragmatique de la connaissance<sup>3</sup>. L'écran n'a pas plus institué l'hypertexte que l'utilisation des index n'a institué le *codex*. Plutôt que la perspective finaliste de Vandendorpe il nous semble plus judicieux de parler d'une co-détermination liée aux usages et à la nature des connaissances dont ils permettent de rendre compte.

<sup>2</sup> voir chapitre premier, point 5.1.1. « Clôture et finitude : un texte à un début et une fin. »

<sup>3</sup> voir le point 7.4.4. « Couplages structurels. » de ce chapitre.

Nous reprenons pour l'exposé de cette partie la triple caractérisation de [Dall'Armellina 00] qui fait de l'écran une « *fenêtre culturelle* », un « *support technique* » et un « *dispositif cognitif* ». Après avoir décrit le fenêtrage culturel mis en place par les écrans qui rythment notre quotidien, nous rappellerons son évolution technique pour dégager un certain nombre de spécificités directement liées à l'hypertexte. Nous verrons enfin, au vu du dispositif cognitif qu'il institue, quelles sont les perspectives d'évolution qui peuvent aujourd'hui être anticipées.

### **1.1. L'écran comme « fenêtre culturelle ».**

Il est une sociologie de l'écran, qui, si elle ne constitue pas notre angle d'approche, ne saurait ici être ignorée : « *Fasciné de surcroît, le co-mutant [= commutant] adhère à son écran comme s'il voulait se fondre dans sa machine : espoir presque réalisé : l'américain passe la quasi-totalité de son temps de conscience éveillé devant des écrans utilitaires ou ludiques.* » [Forget & Polycarpe 97 p.95]. Pourtant, nombreux sont ceux qui comme [Epstein 01] prophétisent que :

« (...) la lecture sur écran (...) ne deviendra peut-être jamais, à mon avis, le mode de distribution majeur des livres en ligne. La perspective la plus probable selon moi, est que la plupart des documents numériques seront imprimés et brochés à la demande dans des points de vente, par des machines – actuellement en prototypes – qui, en l'espace de quelques minutes, produiront à bas prix et à l'unité des exemplaires que l'on ne pourra pas distinguer des livres produits industriellement. »

Cette tendance tient essentiellement aux contraintes de la lecture sur écran (résolution d'affichage, texte reflété ...), contraintes que diverses avancées technologiques s'efforcent de pallier. Ainsi, si les modes de « lecture » sur écran sont légitimement soumis à caution, tous s'accordent à reconnaître que les modes de consultation qu'il autorise sont par contre en pleine expansion, et constituent le cœur de l'architecture numérique qui structure l'édifice de notre réalité quotidienne : « *Telle est aujourd'hui notre architecture : d'immenses écrans sur lesquels se reflètent des atomes, des particules, des molécules en mouvement. Ni une scène publique ou même un espace public mais de gigantesques espaces de circulation, de ventilation et d'éphémères connexions.* » Jean Baudrillard<sup>4</sup>.

### **1.2. L'écran comme « support technique ».**

#### **1.2.1. Historique.**

Du *volumen* au *codex* puis à l'hypertexte, le livre-support a son histoire. A chacune de ses étapes correspond la mise en place de nouveaux modes de pensée<sup>5</sup>. Il en est de même pour l'écran. La mise en perspective historique des mutations de l'écran, établit qu'à l'image du livre et de ses mutations, il est plus qu'un simple support : ses mutations vont en effet permettre de mesurer l'ampleur et le sens de tout un ensemble de paramètres sociaux et de postures cognitives parfois inédites, souvent renouvelées.

<sup>4</sup> cité par [Mathews 93].

<sup>5</sup> voir chapitre premier, point 1 « Le livre. »

acteurs s'exprimant en leur nom propre ou en celui d'une collectivité dans FoRSIC, sont *de facto* situés à la périphérie du système : à chaque fois qu'ils ajoutent une ressource, un besoin, qu'ils s'attribuent une compétence, ou qu'ils consultent les données présentes, l'interface offre la possibilité de « projeter » leur profil, leur compétence, leur besoin ou leur ressource dans l'une des représentations synthétisant celles de l'ensemble du collectif : chaque connaissance, chaque savoir peut ainsi être instantanément resitué et recontextualisé par rapport à l'ensemble des connaissances disponibles. On observe alors des logiques de positionnement très fortes qui renforcent la singularité de certains savoir-faire ou donnent au contraire naissance à des dynamiques d'échange, structurantes pour l'ensemble du collectif.

Cette position à la périphérie, qui facilite et renforce la vue globale et l'observation des dynamiques d'échange à l'œuvre dans le dispositif est par ailleurs complétée par le fait, tout aussi essentiel, que tous les acteurs sont situés à égale distance du cœur du dispositif : chacun peut en effet accéder à l'ensemble des représentations et des données présentes dans le système.

« *cela paraît important de ne pas simplement fragmenter ou décomposer les tâches pour les rendre en elles-mêmes et pour des individus, didactiquement traitables. Toute décomposition de la tâche doit être faite en gardant un œil non pas sur l'utilisateur ou la tâche pris isolément, mais sur le besoin de l'apprenant de situer la tâche décomposée dans le contexte d'une pratique sociale générale. La présence du contexte global offre une chance à l'apprenant de « dérober » ce qu'il ou elle trouve le plus approprié. Il est d'une importance vitale de ne pas fragmenter la périphérie sociale.* ». [Brown & Duguid 92]

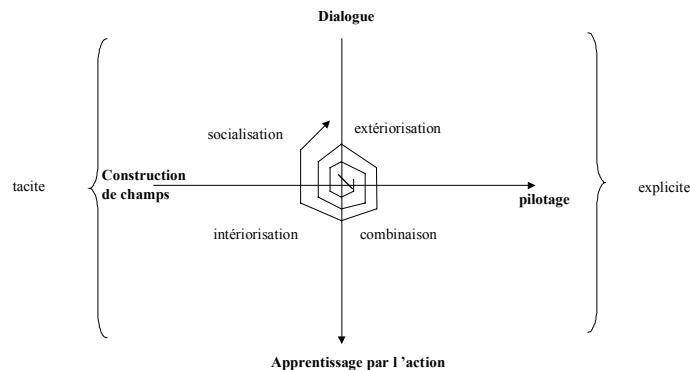
### **6.5. Dimensions.**

FoRSIC, à l'image de l'organisation hypertextuelle qu'il permet de mettre en œuvre et d'évaluer dans un cadre méthodologique – celui de la formation des usagers – et dans un secteur scientifique donné – celui de la recherche d'information – se caractérise par trois dimensions, par trois tendances qui assurent sa cohérence et caractérisent ses spécificités : il est un lieu dans lequel prennent place et s'organisent des logiques de *coopération*, ces logiques sont principalement le fait du sentiment d'appartenance nécessaire à une *communauté de pratiques et d'usages*, elles sont étayées et relayées par des *représentations cartographiques* rendant compte de la topographie des agencements collectifs d'énonciation qui structurent la notion de collectif.

Ce sont ces trois dimensions (coopérative, collective et cartographique) que nous allons maintenant détailler.

#### **6.5.1. Dimension coopérative.**

« *Deux remarques doivent être faites concernant ce processus de montée vers des formes de coopération de plus en plus vastes et intégrées. La première concerne l'accélération du rythme. Chaque nouveau stade de coopération est atteint beaucoup plus rapidement que le précédent. (...) La seconde remarque (...) a trait à l'augmentation de la diversité, ou de la compétition entre différentes formes concurrentes. Plus le stade de*



**Fig. 20 : La conversion de connaissance d'après [Takeuchi & Nonaka 95].**

Afin de clarifier notre exposé, nous nous contentons pour l'instant de décrire les principes théoriques utilisés dans FoRSIC. Nous développerons ensuite la manière dont ces principes nous ont permis d'élaborer nos propres méthodologies et de les mettre en œuvre. Le modèle de conversion des connaissances de [Takeuchi & Nonaka 95] est ainsi à la base des niveaux de formation déterminés pour le champ de la formation à la recherche d'information et a permis d'identifier une série d'invariants correspondant aux « actes documentaires ».

Dans FoRSIC, cette conversion de connaissance est opératoire au niveau des individualités composant le collectif mais également, par une imperceptible variation du niveau d'échelle, à celui de l'ensemble du collectif comme entité homogène.

#### **6.4.4. « Legitimate Peripheral Participation » (LPP).**

Nombre de dispositifs de formation à distance s'efforcent habituellement de réduire au maximum la distance entre l'apprenant (récepteur) et l'enseignant (émetteur), comme si le message, comme si le savoir transmis risquait de se dissoudre ou de s'évaporer dans cette distance. Il nous semble au contraire que la distanciation est nécessaire et qu'elle doit être augmentée – en restant corrélée à l'ensemble des parties intervenant dans le système – pour pouvoir rendre compte de l'environnement complexe dans lequel prennent place les connaissances faisant ou non l'objet d'un enseignement, à la manière d'une catharsis théâtrale qui ne peut se mettre en place qu'une fois la distance au texte établie. L'organisation hypertextuelle, notamment par la topologie des situations de consultation qu'elle met en place, permet d'instituer cette distance en maintenant un arrière plan de signification homogène – dans le cas d'un dispositif donné.

Nous avons dans FoRSIC, mis en œuvre le principe de LPP défini par [Lave & Wenger 91] selon lequel « Les gens ne prenant pas part directement à une activité d'apprentissage particulière apprennent beaucoup de leur positionnement légitime à la périphérie. » [Brown et al. 89] Tous les intervenants, tous les

Pour une analyse détaillée de cette évolution nous rappellerons ses principales étapes, résumées par [Laufer 89 p.207] dans son « *Panorama des techniques et des usages du texte sur écran* » :

« 1- l'écran cinématographique : le texte a fait son apparition sur l'écran, en alternance à l'image, dès les débuts du cinéma. Palliant l'absence de son, les panneaux de texte (...) avaient une fonction narrative.

2- l'écran télévisuel : apporte un considérable élargissement à la place du texte.

3- l'écran informatique : (...) Ce n'est qu'à partir des années soixante que l'on eut recours systématiquement à la visualisation sur écran dans le dialogue avec l'ordinateur.

4- l'écran vidéotex : (...) La page télématique (...) est donc astreinte au format de quarante caractères sur vingt-cinq lignes et à un mode d'organisation, en écrans successifs, que le lecteur consulte sans chevauchement possible d'un écran sur l'autre. Ces contraintes ont engendré une typologie textuelle où l'unité dominante est le paragraphe écran, avec titres, références et boîtes de commande, disposées sur la périphérie de l'écran.

5- l'écran traitement de texte : (...) L'idée est venue d'utiliser conjointement les possibilités de l'ordinateur et celles de l'écran cathodique pour réaliser des machines dont la finalité ne serait plus la manipulation de données - souvent chiffrées -, mais celle du texte en langage naturel.

6- l'écran multi-fenêtré texte-image.

On est donc en présence d'une évolution qui mène de la page-écran purement textuelle à l'écran texte-image, ou, plus exactement, à la mise en page dynamique et intégrée d'écrans à fonctions différentes. »

#### **1.2.2. Spécificités.**

Au-delà de ses propres contraintes de lecture (essentiellement liées à l'affichage), l'écran est par nature soumis à une double contrainte qui donne la mesure de sa spécificité :

« L'écran se définit fondamentalement comme un périmètre clos ; il circonscrit l'information dans les limites d'une surface dont ses bords constituent une clôture. Ce périmètre de validité de l'information confère à l'écran une fonction régulatrice : l'écran est, en ce sens, à la fois une médiation instrumentale (ce qui permet de voir) et une médiation censurante (il admet certaines informations et, métonymiquement, en rejette d'autres). » B. Lamizet<sup>6</sup> cité par [Carré 92 p.75]

##### **1.2.2.1. « Médiation instrumentale » : l'écran protéiforme.**

Cette médiation instrumentale de l'écran est visible sous deux angles : celui de l'espace topologique qui le détermine, et celui lié à la nature de l'information qui « s'inscrit » dans cet espace (écriture/oralité).

Concernant le premier aspect, la médiation de l'écran est une médiation de fait. Il n'est d'ailleurs guère difficile de mettre au jour l'héritage de la page-papier dans la détermination de la page-écran, comme le fait [Lévy 90 p.55] « La largeur de l'écran [aux débuts du Mac] fut calculée à partir du format standard d'une feuille de papier. » Actuellement, « un écran 21 pouces » contient « la double page d'un livre grand format. » [Laufer & Scavetta 92 p.53] L'écran est donc en premier lieu une instrumentalisation de la page traditionnelle selon une dynamique d'héritage maintes fois repérée dans ce travail pour plusieurs des aspects de l'organisation hypertextuelle. C'est l'infrastructure technique des réseaux et des connections qu'ils autorisent qui est ici seule source de spécificité. [Lévy 97] « On parle de page dans les deux cas [feuille de papier et page-écran], mais la première est un pagus, un champ borné, approprié, semé de signes enracinés, l'autre est une unité de flux, soumise aux contraintes du débit dans les réseaux. »

<sup>6</sup> B. Lamizet, « Une nouvelle communication : l'écran entre le texte et l'image. », *Quaderni*, n°8, Automne 1988.



Le second aspect nous semble par contre ne plus s'inscrire dans un héritage mais le renouveler. Pour [Lévy 90 p.10] « *La succession de l'oralité, de l'écriture et de l'informatique comme modes fondamentaux de gestion sociale de la connaissance ne s'opère pas par simple substitution, mais plutôt par complexification et déplacement des centres de gravité.* » L'écran par lequel on accède à l'hypertexte subsume les trois : son support est clairement informatique, il continue de véhiculer essentiellement des textes (donc de l'écrit) et il reprend à son compte des codes de communication habituellement associés à l'oralité (interfaces dialogiques, avatars ...). L'écran est le premier de ces centres de gravité de l'édifice numérique : il tient ensemble les deux dimensions de la communication humaine, l'oralité et l'écriture. Plus que cela, il établit entre elles des logiques inédites : il autorise une lecture/écriture de l'oralité (sites de « chat », « chatrooms »<sup>7</sup>) et oralise des processus d'écriture traditionnels (apparition des émoticônes ou « smileys » par exemple).

#### 1.2.2.2. « Médiation censurante » : l'écran procustéen.

Vu sous un autre angle, l'écran est une « médiation censurante ». La logique inédite qui se met ici en place ne se mesure plus à l'aune des rapports renouvelés entre écriture et oralité, mais bien dans celle d'un rapport individuel à notre propre corps.

*« L'écran de nos images, l'écran interactif, l'écran télématique sont à la fois trop proches et trop lointains : trop proches pour être vrais (pour avoir l'intensité dramatique d'une scène), trop loin pour être faux (pour avoir la distance complice de l'artifice). Ils créent de la sorte une dimension qui n'est plus exactement humaine, une dimension excentrique qui correspond à une dépoliarisation de l'espace et à une indistinction des figures du corps. » [Baudrillard 88 p.19]*

Comme a tenté de l'établir notre second chapitre<sup>8</sup>, le réseau est la seule possibilité d'existence pour le corps social. Les individualités qui le composent ne disposent comme miroir, que de l'image reflétée que leur renvoie l'écran par le biais duquel elles « entrent » sur le réseau. La posture lectorale est alors doublement inédite : dans ses modalités elle devient une posture perceptive directement réglée par la ou les navigations possibles dans cet environnement affiché sur écran (avatars, réalité virtuelle, etc ...) Quand elle demeure au premier plan une activité de type lectorale, c'est là encore selon des modalités inédites.

*« (...) l'écran apparaît comme le point d'aboutissement du mouvement qui a séparé le texte du corps. (...) La nouvelle position de lecture (...) est radicalement originale : elle assemble, et d'une manière qui serait encore à étudier, des techniques, des postures, des possibilités qui, dans la longue histoire de la transmission de l'écrit, restaient séparées. » [Chartier 97 p.13]*

Nous avons mis au jour certaines de ces possibilités dans notre annexe 4 « Stratégie des interfaces ».

<sup>7</sup> Bavardage. Sites sur lesquels des internautes se retrouvent pour parler « en temps réel ».

<sup>8</sup> Point 1.4. « Logiques de l'adéquation. »

incarnée qui appartient à chaque individu) et des systèmes de connaissance socialement distribués, qu'apparaissent des propriétés cognitives non prévisibles.

Dans cette lignée nous nous proposons d'étudier la communication et la coopération à travers la propagation d'états représentationnels de connaissances internes (propres à un individu) et externes (qui appartiennent au groupe et aux supports d'information qu'ils utilisent) pour supplanter le déterminisme émetteur - récepteur de la communication dont nous avons par ailleurs<sup>71</sup> pointé les nouveaux modes de déploiement dans le cadre de la carte énonciative hypertextuelle.

#### 6.4.3. Le modèle de la conversion des connaissances.

[Takeuchi & Nonaka 95] proposent une modélisation opératoire retraçant le cheminement qui mène de la connaissance tacite à la connaissance explicite, sous la forme de quatre modes de conversion en interaction permanente. Ces quatre modes sont les suivants :

- **la socialisation** : elle concerne les procédures. Ce type de connaissance peut s'acquérir directement en observant quelqu'un (observation, imitation, mise en pratique). Mais pour que cette connaissance soit efficace, il faut la confronter aux autres types de connaissance, la finaliser, l'incarner dans un travail à faire, une tâche à réaliser et la mettre en perspective sociale. Elle privilégie l'échange de savoir-faire et d'informations non-structurées ;
- **l'extériorisation**, ou passage du tacite à l'explicite est un moment d'articulation vers le niveau conceptuel et s'exprime à travers des analogies, des métaphores. Utilisant l'induction et la déduction, ce moment fait émerger la nouveauté. Cette expression de la réalité se fait en concepts plus clairs, plus adaptés. C'est un moyen pour explorer, problématiser et conceptualiser une question, construire son environnement sémantique en la situant dans un ensemble ;
- **la combinaison** (de l'implicite à l'explicite) : c'est le moment de la transformation des concepts en système de connaissance où interviennent et se croisent plusieurs ensembles de connaissances explicites. Cette ré-ingénierie conceptuelle est atteinte au moyen de nouveaux tris, d'ajouts, de mises à jour. La connaissance nouvelle vient s'inscrire dans une hiérarchie de savoirs, de façon systémique : elle trouve sa place dans un ensemble déjà structuré ;
- **l'intériorisation** (de l'explicite au tacite) est l'incarnation de la connaissance explicite sous une autre forme. L'individu ayant acquis de nouveaux savoirs peut les incarner dans de nouveaux travaux. Il s'approprie la connaissance en cours de développement, y compris la connaissance d'autrui et l'incarne dans un produit. Ce niveau de connaissance vient s'agréger aux croyances de l'utilisateur, vient enrichir sa vision du monde à un niveau tacite et permet de produire de nouveaux objets de connaissance à partir de la transformation des informations nouvellement acquises et métabolisées ;

<sup>71</sup> voir notamment le point 4 « Emergence de nouvelles subjectivités. » du chapitre premier.

### **6.3.2. Systémique.**

« A la différence de l'approche analytique, l'approche systémique englobe la totalité des éléments du système étudié, ainsi que leurs interactions et leurs interdépendances. (...) Il s'agit de dégager des invariants, c'est-à-dire des principes généraux, structuraux et fonctionnels, pouvant s'appliquer aussi bien à un système qu'à un autre. » [Rosnay 75 p.92]

Nous nous situons dans une perspective épistémologique issue de la systémique mais qui la dépasse, considérant que l'émergence est augmentée par la coopération. La signification émerge d'agencements collectifs d'énonciation (distribution sociale de la connaissance) qui coopèrent et communiquent.

### **6.4.Principes.**

Ce cadre générique étant – rapidement – posé, nous présentons maintenant les approches mises en œuvre dans FoRSIC.

#### **6.4.1. Couplage structurel.**

Dans FoRSIC, le comportement global du système n'est pas le fruit d'un contrôle exercé par le dispositif central<sup>70</sup>. En termes de validation par exemple, FoRSIC ne dispose d'aucun comité éditorial amené à se prononcer sur la validité des ressources de formation qui y sont déposées : le seul processus de validation ayant cours est celui d'une validation par l'usage collectif, qu'il s'agisse des ressources, des compétences, ou de tout autre élément entrant dans le système.

Ainsi, le comportement global émerge de l'organisation issue des connexions internes et des couplages structurels avec l'environnement. Le couplage structurel est un processus de comportement dynamique non figé, lié au sujet et qui permet « *de faire émerger de la signification sur un arrière-plan de compréhension* » [Varela et al. 93 p.149]. Cette perspective théorique s'applique au système de production de connaissance qu'est l'individu confronté à son environnement informationnel et donne un cadre de travail pour décrire la dynamique des relations entre le sujet et son environnement et la dynamique de la production des connaissances.

#### **6.4.2. Cognition distribuée.**

[Hutchins 95] fait de l'homme le site de l'information et propose le concept de cognition distribuée dans le cadre de l'étude de tâches réelles complexes. Pour lui, la communication n'est pas un simple processus de transfert de connaissances d'un agent à un autre mais renvoie à la création d'une nouvelle connaissance. C'est à partir de ce qu'il appelle « *locus of knowledge* » (site de la connaissance mémorisée,

<sup>70</sup> cette absence de contrôle central est une propriété essentielle que partagent l'hypertexte, le réseau et le rhizome (voir le point 1.4 « Logiques de l'adéquation. » du chapitre second.

### **1.3. Quel avenir pour ce dispositif cognitif ?**

Plusieurs critères qui engagent l'avenir de l'écran pour les prochaines années peuvent être isolés.

#### **1.3.1. Nature de l'écran : entre surface et interface.**

« (...) la modification des notions habituelles d'espace et de temps a étendu le principe de la relativité à l'ancienne notion de superficie, de surface, aboutissant au remplacement de celle-ci par la notion d'interface, notion trop mésestimée qui établit pourtant la superficie dans sa relation médiatique, et qui fait de cette dernière un 'face à face', un 'vis à vis'. » [Virilio 90 p.71]

##### **1.3.1.1. Surface écran.**

Si l'écran est une surface, celle-ci – comme celle du papier qui l'a précédé – se caractérise au premier chef par ses possibilités d'inscription. Une inscription qui dispose de ses propres règles d'écriture (multimédia, gestion des liens, rupture avec la linéarité) et, encore à titre expérimental, de ses propres outils<sup>9</sup>. L'espace d'inscription des signes que circonscrit cette surface devient un espace actif :

« J.P. Balpe note que, dans les dispositifs hypertextuels, « l'espace sur lequel les signes s'inscrivent n'est plus un espace inactif, mais un espace actif enfermant une infinité d'espaces virtuels. Un signe sur un écran est un signe au même titre que sur une page quelconque, mais il est, également, signe toujours possible d'une multiplicité d'autres niveaux de signes. Le signe n'est plus seulement signe d'un sens, mais signe de transformations virtuelles. » [Lévy 91 p.128]

##### **1.3.1.2. Interface écran.**

« La surface du grand écran s'est mue en interface, petit écran médiateur de la télévirtualité. » [Veillon 97]

A l'inverse, si l'écran est d'abord une interface, celle-ci perd toute possibilité d'inscription<sup>10</sup>. Ce qu'elle autorise n'est plus une lecture « sur écran » mais une lecture « de l'écran » : « *La lecture d'un écran est tout à fait différente de celle du regard. C'est une exploration digitale, où l'œil circule selon une ligne brisée incessante. Le rapport à l'interlocuteur dans la communication, au savoir dans l'information est du même ordre : tactile et exploratoire.* » [Baudrillard 88 p.19]

#### **1.3.2. Dimension de l'écran.**

« Au vingt et unième siècle, celui qui contrôlera l'écran, contrôlera la conscience. » Timothy Leary. Cité par [Virilio 90 p.113]

Ce choix qui reste encore à faire entre la surface et l'interface – et qui se fera probablement sous la forme de modalités conjointes – est un écho de celui qui se pose à l'échelle de l'organisation hypertextuelle

<sup>9</sup> l'E-ink (encre électronique) : voir le site <http://www.eink.com>.

<sup>10</sup> à propos des « écrits d'écran » voir notamment les travaux de Jeanneret et Souchier. [Jeanneret & Souchier 02] en offre une synthèse.

entre la part accordée au discours et celle accordée au dispositif<sup>11</sup>. Le critère déterminant pour l'écran sera probablement celui de sa stature, de sa dimension.

Il est en effet possible d'établir un rapport direct entre la nature des écrans et la taille qu'ils occupent et sous laquelle ils s'offrent au regard. L'échelle dont il est ici question se déploie du macro (« Jumbotrons » dont parle [Virilio 90 p.11] ou écrans des stades olympiques) au micro (les terminaux actuellement les plus courants ont comme point commun la miniaturisation de leurs écrans, qu'il s'agisse d'ordinateurs, de téléphones portables – technologie WAP et palms<sup>12</sup> notamment – de casques de réalité virtuelle ...).

Au niveau macro l'écran est avant tout une surface sur laquelle il est ou non possible d'agir dynamiquement. Au niveau micro, l'écran est avant tout interface. Dans l'un comme dans l'autre cas [Bougnoux 93 p.788] « *L'espace de rencontre de l'écran est, en effet, défini par la configuration du système technique* ».

Les voies de recherche actuellement les plus développées sont celles de l'écran-interface. Pour [Negroponte 95 p.119] « *Le défi de la prochaine décennie n'est pas de mettre au point de plus grands écrans (...) [mais de] fabriquer des ordinateurs qui vous connaissent, n'ignorent rien de vos besoins et comprennent le langage verbal et gestuel* ». Pour autant (les techniques holographiques et celles d'immersion en réalité virtuelle restant encore souvent de nature expérimentale), du point de vue qui nous occupe, une voie moyenne est à définir et à creuser : celle qui en augmentant la taille (surface) des écrans dans les limites d'une perception individualisée et non collective, permettra de décliner des modes d'interaction (interface) que l'organisation hypertextuelle rend aujourd'hui nécessaires (transclusion et versioning notamment).

Les individualités composant le corps social constitué sur les réseaux pourront peut être alors disposer, à l'échelle de leur corps singulier, d'un accès à l'une des dernières frontières du cyberspace : « *Timothy Leary parle de « Psybernétique ».* Selon lui « *la prochaine frontière est notre propre cerveau, le 'neuroworld' ».* » [Donguy 99]

<sup>11</sup> voir le point 8.5. « La forme des genres. » du chapitre premier.

<sup>12</sup> la technologie WAP est utilisée pour permettre à des téléphones portables de se connecter à Internet. Les Palms sont des agendas électroniques.

### **6.3. Fondements théoriques.**

Nous voulons ici succinctement présenter les deux grandes approches auxquelles nous nous référons dans FoRSIC, le connexionnisme et la systémique. Il ne s'agit évidemment pas d'un historique de ces champs théoriques complexes ni d'une discussion épistémologique sur leurs fondements, leur portée et leurs limites, mais simplement d'un recadrage pour situer les axes de recherche qui ont présidé à la mise en œuvre de ce projet.

#### **6.3.1. Connexionnisme.**

« *De même que nous ne pouvons absolument pas concevoir des objets spatiaux en dehors de l'espace ni des objets temporels en dehors du temps, nous ne pouvons imaginer aucun objet en dehors de la possibilité de sa connexion avec d'autres objets. Si je puis concevoir l'objet dans le contexte de l'état de choses, je ne puis le concevoir en dehors de la possibilité de ce contexte.* » [Wittgenstein 61 p.30]

Pour définir le connexionnisme, [Varela et al. 93 p.34] expliquent :

« *Ce nom dérive de l'idée que de nombreuses tâches cognitives (à titre d'exemple, la vision et la mémoire) semblent être effectuées de manière optimale par des systèmes consistant en un grand nombre de composants simples qui, quand ils sont connectés selon des règles appropriées, donnent lieu à un comportement global correspondant à la tâche désirée. (...) Par opposition au cognitivisme, les modèles connexionnistes remplacent généralement le traitement localisé, symbolique par des opérations distribuées, c'est-à-dire des opérations qui s'étendent au réseau entier des composants ; c'est ainsi qu'ils produisent l'émergence de propriétés globales résistant à une dysfonction locale.* »

FoRSIC s'inscrit dans ce cadre théorique. Le domaine cognitif sur lequel il se focalise est celui de la formation, c'est-à-dire de la transmission de la connaissance. Comme nous le préciserons dans la partie consacrée aux ontologies et à leurs couplages, et comme cela a déjà été décrit pour l'architecture fonctionnelle et les quatre bases sur lesquelles elle repose, chaque élément du dispositif est connecté aux autres selon des règles spécifiques. L'optimisation de la tâche assignée au départ (la formation des usagers par la mutualisation et le partage de connaissances), est visible dans le comportement global du système, et a des effets rétroactifs sur chacune de ses composantes (comportement du collectif de formateurs, évolution du modèle de domaine et du référentiel de compétences, évolution de la place de chacun au sein du collectif, logiques d'échange, de partage, etc.).

Nous pensons dans le même temps qu'un modèle connexionniste simple n'est pas suffisant. Comme nous l'avons souligné à propos de l'organisation hypertextuelle, il ne suffit pas d'établir des connexions entre différents éléments pour qu'une dynamique d'échange et de transmission de savoir soit opératoire et surtout interprétable par ceux auxquels elle est destinée. Ce modèle connexionniste initial doit pouvoir être amorcé et complété par un certain nombre de modèles formels qu'il lui appartiendra alors de dépasser ou de conserver.

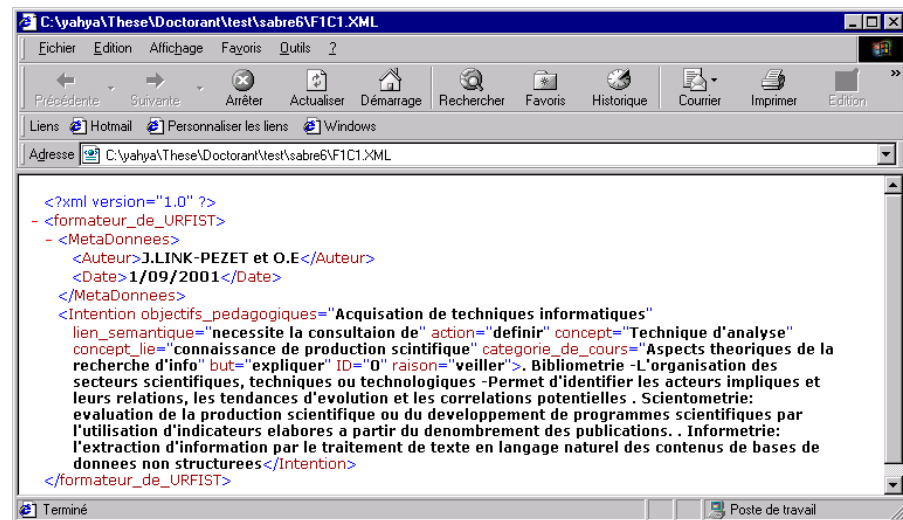


Fig. 18 : SABRE : un exemple de ressource XML produite.

## 2. Bibliothèque.

« Symboliquement, le monde antique finit avec la destruction de la bibliothèque d'Alexandrie ; symboliquement, le XX<sup>ème</sup> siècle s'achève avec la reconstruction de la bibliothèque de Sarajevo. » Alberto Manguel, *Dans la forêt du miroir*, Actes Sud. Cité par [Bermond 00]

Nous voulons ici, au-delà de l'inventaire des modalités d'existence des bibliothèques (liées à l'archivage, à la conservation et à l'accès qu'elles permettent, aux supports dont elles ont la charge), questionner la symbolique qu'elles véhiculent. La bibliothèque, à son tour, pose de manière obsédante la question du lieu : lieu de consécration de l'objet-livre, lieu de consultation et d'accès au(x) savoir(s), lieu qui organise la connaissance pour la distribuer selon les infrastructures technologiques disponibles à une époque donnée. Classique, électronique puis virtuelle, la bibliothèque est le premier de ces lieux collectifs dans lesquels l'organisation hypertextuelle se donne à voir.

### 2.1. Classique, électronique, virtuelle ...

Les mutations des bibliothèques sont de deux ordres : celles liées à l'organisation des savoirs qu'elles permettent de conserver, et celles des traces, des supports sur lesquels sont inscrits ces savoirs. C'est cette double articulation qui permet de distinguer la bibliothèque classique, de la bibliothèque électronique (« l'informatisation de la bibliothèque classique et la numérisation des textes ») et de la bibliothèque virtuelle (« la bibliothèque électronique plus le réseau, plus l'appropriation individuelle. » [Giffard 96]).

Le passage à la bibliothèque électronique relève d'un changement de support quand la bibliothèque virtuelle procède d'une dynamique plus générale liée aux nouvelles formes d'organisation du savoir sur les réseaux. Ambitions intellectuelles et contraintes technologiques sont par ailleurs, et depuis toujours, les deux faces d'une même réalité bibliothéconomique : dès 1576, « *Alde Manuce réussit à fusionner le programme intellectuel de l'édition humaniste et le programme technologique de la typographie.* » [Giffard 96]

L'avènement du numérique et de l'organisation hypertextuelle dont il procède, pose cependant de nouveaux problèmes.

Premièrement, les bibliothèques virtuelles, tout en maintenant leur vocation de diffusion/conservation, « *reçoivent de l'information et non seulement des œuvres* » [Bougnoux 93 p.658], ce qui, deuxièmement, « (...) conduit à poser en d'autres termes la dialectique du « périssable » et du « conservable ». Il nous permettra peut-être de constituer et de transmettre un « patrimoine du temporaire ». » [BPI 01]. Là encore, le problème du versioning demeure l'un des principaux champs d'investigation.

Enfin, troisièmement, « *La technologie permettant de délivrer d'immenses quantités de données ne délivre pas simultanément une raison légitime d'accumuler ces informations.* » [Friedlander 95 p.75]<sup>13</sup> Et

<sup>13</sup> sauf à s'inscrire dans la voie futuriste de la bibliothèque universelle comme en témoignent les dialogues de l'épisode 2 de la trilogie « Star Wars » entre Jocasta Nu (cyber-bibliothécaire) et Obi-Wan (usager) :

- ObiWan : (...) peut-être les archives sont-elles incomplètes. Jocasta Nu : Les archives sont vastes et parfaitement sûres mon jeune Jedi. Il est une chose dont vous pouvez être absolument sûr : si une donnée n'apparaît pas dans nos enregistrements, elle n'existe pas. » Script téléchargeable à l'adresse <http://www.sithclan.net/episode2/scriptEP2.doc>, consulté le 03/06/2002.

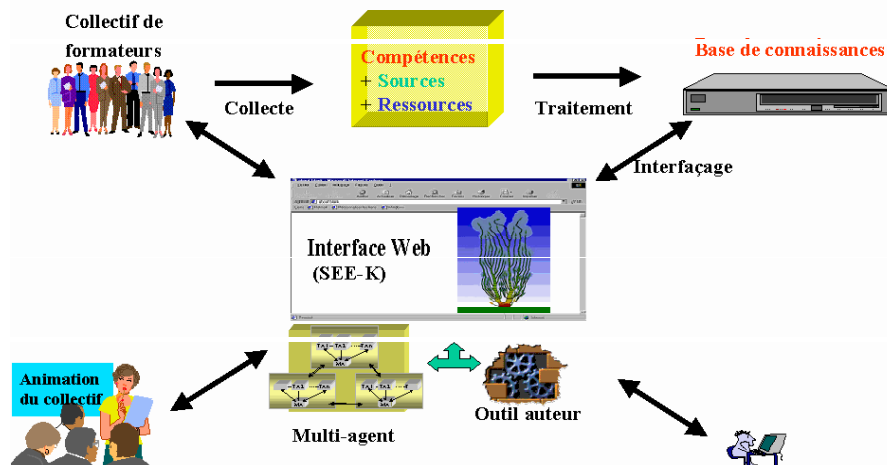


Fig. 19 : Architecture fonctionnelle de FoRSIC.

avant même de songer à trouver une justification à la conservation de ce « *patrimoine du temporaire* », il faut organiser et homogénéiser les systèmes de classification permettant d'y accéder (Dublin Core, Métadonnées, TEI<sup>14</sup> ...).

## 2.2. Pour un nouvel encyclopédisme.

### 2.2.1. Encyclopédisme savant.

Confronté à la masse d'information potentiellement recevable par les bibliothèques, aux difficultés à organiser cette masse de données et aux choix d'archivage qui sont en permanence opérés et remaniés, c'est la perspective encyclopédiste, demeurée intacte depuis la renaissance, qui se trouve remise en question.

*« Au catalogue actuel, ne faut-il pas substituer ou au moins adjoindre un système de cartographie intelligente qui permette de restituer chaque donnée dans son « arbre de savoir » ? (...) On observera que c'est aussi le chemin obligé d'une redéfinition de l'encyclopédisme, question qui est aujourd'hui massivement posée. »* [BPI 01]

L'encyclopédisme n'est en effet valide que tant qu'il peut prétendre à une certaine stabilité, au maintien de certaines linéarités :

- stabilité temporelle qui permet de maintenir un décalage minimal entre les avancées du savoir et les encyclopédies en rendant compte, dans une perspective cumulative ;
- linéarités croisées, transversalités permettant d'offrir des points d'entrée communs à un ensemble de domaines, de techniques.

Or, comme le note [Varet 56 p.69] et ce avant même les premières entreprises encyclopédistes :

*« Le savoir affecte forcément une forme circulaire : c'est en effet la seule manière de se représenter un ensemble de données diverses tel que chacune renvoie à toutes les autres et ait perspective sur toutes les autres. (...) Ce savoir n'est pas simplement cumulatif (...) mais circulaire parce qu'il y a une circulation du savoir d'un point quelconque à tout autre point possible. Sans doute cette circulation se fait elle le long de certains axes perspectifs qui seront par habitude plus fréquentés que d'autres à l'intérieur du tout, mais dont la commodité ne tient jamais finalement qu'à un état momentané du savoir, à un équilibre météorologique métastable. »*

S'il est effectivement question de linéarités comme autant « d'axes perceptifs momentanés », elles ne sont que des commodités de l'analyse et ne témoignent en aucun cas de la nature des objets dont elles peuvent momentanément permettre de rendre compte.

Nous avons montré, notamment dans notre second chapitre, qu'à l'inverse de ce qu'écrit [Zeldin 01] pour qui « *L'Internet en est encore à l'âge de Bouvard et Pécuchet dont l'ambition était de copier toutes les connaissances.* », nombre de manières d'organiser ces connaissances sont déjà opératoires sur l'ensemble des réseaux de communication et d'échange de savoirs, et que de nombreux et significatifs efforts sont mis en place pour harmoniser et relier ces différentes approches, tendant à l'instauration d'un encyclopédisme d'usage(s).

<sup>14</sup> Le Dublin Core, les métadonnées, et la Text Encoding Initiative (TEI) sont des projets permettant d'harmoniser et de structurer l'indexation des données sur le réseau en balisant les documents à l'aide de titres, de mots-clés, de descriptions bibliographiques, etc. La plupart de ces spécifications sont accessible depuis le site du consortium W3 (<http://www.w3c.org>)

SABRE est ainsi un outil qui permet de produire des ressources de formation individualisées et modulables en autant de versions qu'il peut exister de besoins différents chez les apprenants à qui sont destinées ces ressources. Il offre en outre l'une des toutes premières implémentations concrètes des recommandations prônées par le web sémantique : la possibilité d'effectuer une recherche intelligente dans un corpus de ressource, recherche ne reposant plus uniquement sur l'analyse lexicale assortie d'un croisement (matching), mais une recherche par niveau, par intention, par concept, etc. A partir de cette recherche initiale, il offre également la possibilité d'atteindre – grâce aux liens utilisés dans le réseau sémantique – les ressources les plus pertinentes offrant les pré-requis nécessaires à la compréhension de la ressource trouvée, celles faisant appel aux mêmes concepts, celles appartenant à la même catégorie de cours, etc.

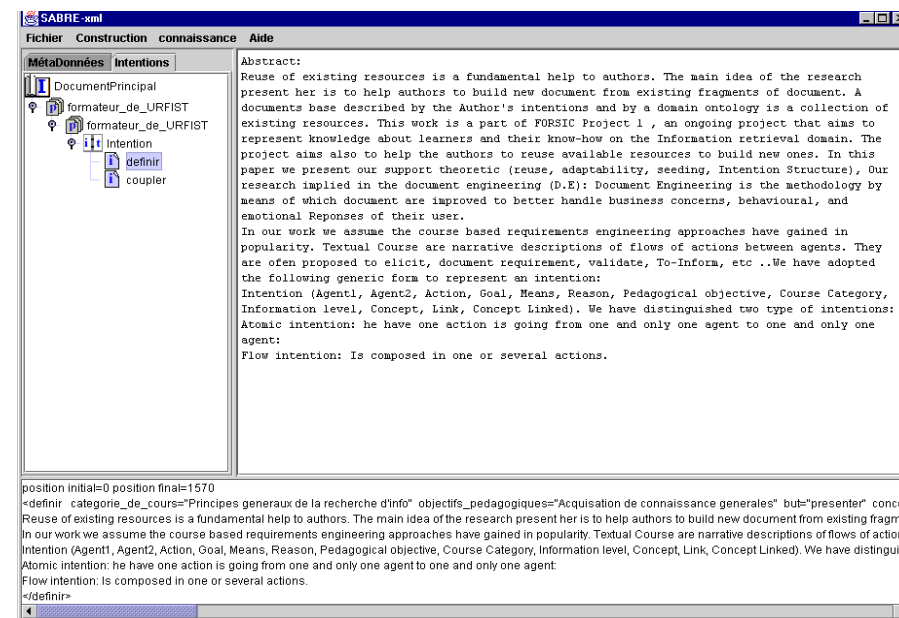


Fig. 17 : Copie d'écran du dispositif SABRE.

### 6.2.2. Un outil-auteur pour la ré-ingénierie documentaire.

Sur la base du modèle de domaine de la recherche d'information que nous avons établi et qui est décrit plus bas, l'outil-auteur SABRE<sup>68</sup> fonctionne sur un réseau sémantique reprenant les termes présents dans ce modèle de domaine reliés par deux relations : « nécessite la connaissance de » (notion de pré-requis) et « est utile à la connaissance de » (notion d'aide à la compréhension). Ces deux relations peuvent être établies entre deux concepts du modèle de domaine mais également entre deux documents ou deux parties de documents présents dans SABRE, ou enfin entre un concept du modèle de domaine et un document (ou partie de document) présent dans SABRE.

SABRE s'articule sur la base des ressources de formation présentes dans FoRSIC. Il se présente sous la forme d'un module indépendant, offrant la possibilité de créer une ressource pédagogique à partir d'une ou de plusieurs ressources déjà présentes, ou bien en commençant à zéro. Les documents sont entrés au format texte ou XML et, en sortie, ils sont tous des documents XML.

L'interface de SABRE offre la possibilité d'associer aux ressources de formation des indications concernant :

- l'intention pédagogique du créateur de la ressource (possibilité d'associer plusieurs intentions à l'ensemble d'une ressource ou à des parties distinctes de la même ressource) ;
- les niveaux de formation mobilisés (explicités plus bas dans ce travail) (possibilité d'associer plusieurs niveaux à l'ensemble d'une ressource ou à des parties distinctes de la même ressource) ;
- la « catégorie » du cours ainsi produit (en corrélation avec le niveau de formation visé) ;
- l'objectif pédagogique du formateur (possibilité d'associer plusieurs objectifs à l'ensemble d'une ressource ou à des parties distinctes de la même ressource) ;
- un ou plusieurs concepts présents dans le modèle de domaine, à un niveau de granularité qui s'arrête au paragraphe ;
- des métadonnées plus « classiques » comme le nom de l'auteur ou la date de création.

SABRE dispose d'une vue arborescente qui permet à l'utilisateur d'observer l'attribution et le marquage synchrone de ces informations, au fur et à mesure de la création de la ressource (ou de sa re-composition à partir de ressources existantes). Dans le cas où l'auteur réutilise des ressources ou des parties de ressources déjà présentes dans SABRE, c'est-à-dire déjà « marquées » à l'aide des entrées que nous venons d'évoquer, il est alors libre de conserver ce marquage en l'état ou de le modifier, donnant ainsi naissance à une version différente du même document<sup>69</sup>.

En sortie de l'outil-auteur SABRE, nous disposons donc de ressources XML dont le marquage reprend cet ensemble d'éléments contextuels qui sont essentiels pour tout document de formation dans un domaine donné (c'est-à-dire dans lequel peuvent être définis des niveaux, des objectifs, etc.)

<sup>68</sup> SABRE : Système Auteur Basé sur la Réutilisation.

<sup>69</sup> En ce sens, SABRE fonctionne sur les principes du versioning.

### 2.2.2. Encyclopédisme d'usage.

La perspective historique dressée par Goody et reprise par Parrochia nous permettra de comprendre la nature du changement qui affecte l'idée encyclopédiste, l'encyclopédisme savant cédant la place à un encyclopédisme d'usage qui s'avère seul pertinent pour rendre compte des modes de transmission et d'organisation de la connaissance sur les réseaux.

Le premier type de classification fut celui introduit par l'usage des listes. « [les listes] *permirent non seulement de fixer en l'état un savoir mais aussi de poser des problèmes de classification au point d'atteindre les limites extrêmes d'un certain type de compréhension du monde.* » [Goody 79 p.169]. Une fois cette limite atteinte, d'autres types de classification furent proposés et prirent place en complémentarité des premières :

*« Historiquement, d'abord on énumère (...), ensuite on rassemble (...), enfin on classe et l'on ordonne (...). Dès lors, trois grands types de structuration du savoir se font jour : le séquentiel linéaire (la liste), le central-circulaire (l'organisation systématique – encyclopédique), enfin le relationnel-réticulé (la base ou banque de données, le néo-fichier). »* [Parrochia 93 p.181]

Ainsi la bibliothèque, quelle que soit sa dénomination (numérique, électronique ...) ne change pas : elle propose et dispose de ces trois invariants, s'inscrit pleinement dans leur cycle. Ce qui change c'est « l'accès » compris comme synthèse du niveau de connaissance et de structuration mentale (styles cognitifs) de l'utilisateur. Quand les formes traditionnelles de la bibliothèque supposaient un accès par l'un des trois modes (procéduraux) présentés ci-dessus pour pouvoir accéder à la connaissance sous forme déclarative, la dynamique s'inverse doublement avec la bibliothèque virtuelle.

Un nouveau scénario voit le jour : l'utilisateur accède (essentiellement grâce à du « *relationnel réticulé* » – voie privilégiée mais non-exclusive) à des connaissances déclaratives (données) à partir desquelles il va tenter de comprendre, d'inférer comment il peut avoir accès à d'autres données liées et considérées comme les plus pertinentes de son point de vue (et non plus de celui des classifications établies *a priori*). L'accès aux documents n'est plus subordonné aux modes de classement et d'organisation. Ce sont les modes de classement et d'organisation qui sont inférés de l'accès au document et de l'analyse de son contenu.

On parlera d'un encyclopédisme d'usage à partir du moment où ces parcours de recherche, d'accès et de consultation sont récupérés et réinjectés dans l'organisation de la bibliothèque virtuelle (Internet en est évidemment le premier exemple) pour organiser, à l'aune de ces parcours, les nouvelles données devant être classifiées<sup>15</sup>.

<sup>15</sup> certains algorithmes de moteurs de recherche déjà mentionnés fonctionnent sur ce principe (google, clever) ; nombre de systèmes établissant des métadonnées se fondent également sur des études de profiling et plus généralement sur des études comportementales des usagers en recherche d'information.

### 2.2.3. De la bibliothèque-mémoire à la bibliothèque neuronale.

« Il y a une mémoire neuronale. Il y a, de nos jours, une mémoire informatique. Dans le deux cas, il y a inscription. »  
[Link-Pezet 89 p.12]

L'organisation, les modes de classification qui émergent ainsi de la prise en compte simultanée des parcours associatifs individualisés d'accès à l'information, augure d'une nouvelle étape dans les processus complexes de mutation qui travaillent les bibliothèques. Conformément aux pistes dégagées dans notre premier chapitre<sup>16</sup>, les conditions semblent réunies pour que s'opère un nouveau transfert de compétences depuis les capacités mnésiques individuelles vers celles, collectives, sur lesquelles s'érigent les lieux de conservation et de diffusion des savoirs.

Comme le note [Atlan 95] à propos des premières :

« On a affaire à une mémoire qui est, comme on dit, « a content address » : l'adresse est déterminée par le contenu, et non plus, comme précédemment par un lieu. (...) Ce qu'on va chercher n'est pas défini par le lieu où on l'a mis (...), mais par le contenu même de ce qu'on va chercher. On pourrait objecter : « Cela veut dire qu'il faut savoir à l'avance ce que je veux me rappeler. » Non ! Il suffit que je sache un tout petit morceau, plus ou moins grand, de ce que je veux me rappeler pour déclencher le rappel de la totalité. Cela ressemble plus à notre façon de nous souvenir ... »

On retrouve ici le précepte fondateur de [Bush 45] : « L'esprit humain fonctionne par associations ». La bibliothèque du « *Neuroworld* » dont parle Leary est désormais à portée de connaissance : ses modes de classement et d'organisation sont ceux qui, indépendamment du support permettent de « déclencher le rappel de la totalité ». Avant d'en arriver au stockage systématique et universel d'unités mémoires toutes interconnectées, les réalités informatiques – ou en passe de le devenir – que constituent déjà la transclusion, le versioning ou certains types de navigation, les recherches menées sur les réseaux neuronaux et synaptiques et leurs applications technologiques, le formidable déploiement de langages universels de balisage, l'essor de l'ingénierie des connaissances et ses implications dans le domaine de la ré-ingénierie documentaire<sup>17</sup>, le web sémantique et la place qu'il accorde à la construction d'ontologies<sup>18</sup> en sont autant d'indices prometteurs.

### 2.3. La bibliothèque sans livres : pour une archivistique topologique.

« (...) mon objet n'est pas le langage mais l'archive, c'est-à-dire l'existence accumulée des discours. » [Foucault 94 p.595]

Le principal argument, tant philosophique que technique qui est développé comme horizon limitatif des mutations possibles de la bibliothèque actuelle, est celui dont fait état [Lepers 97 p.315] :

« L'idée structuraliste d'une universalité des moules de la pensée humaine serait évidemment merveilleuse pour la constitution des hypertextes planétaires ; en réalité, cela relève plutôt de la fable, ou d'un mythe local, reflétant le fantasme universaliste de la rationalité occidentale. Il n'est pas question, bien entendu, de contester les bénéfices évidents de ce fantasme, mais de trouver les moyens d'une autre négociation entre les rationalités locales et la raison universelle. »

<sup>16</sup> point 4.3. « De l'auteur au lecteur. »

<sup>17</sup> voir le point 6.2.2. de ce chapitre « Un outil-auteur pour la ré-ingénierie documentaire. »

<sup>18</sup> voir le point 7 de ce chapitre.

un certain nombre de renseignements administratifs (état-civil, formation initiale), à décrire son environnement de formation (personnes-ressources, publications, etc.) et à « déposer » un certain nombre de besoins (matériels, techniques, de formation ...). Chaque membre du collectif dispose également de la possibilité, via le référentiel sur lequel nous reviendrons plus tard, d'associer un nombre illimité de compétences à sa fiche personnelle, aux ressources de formation qu'il dépose et aux besoins qu'il exprime. A chaque fois, ces compétences peuvent être assorties d'un commentaire en texte libre permettant de les préciser ou de les recontextualiser.

L'outil SEE-K permet ainsi d'organiser trois bases différentes (base formateur, base ressource et base besoin) autour d'une base centrale (base compétence contenant le référentiel) qui permet d'indexer les trois autres, chaque information contenue dans ces bases pouvant être décrite par l'association à une ou plusieurs compétences.

Cet outil produit ensuite des cartographies dynamiques et synchrones de l'ensemble des informations contenues dans les quatre bases. Ces cartographies reposent sur la « philosophie » des arbres de connaissances<sup>67</sup>.

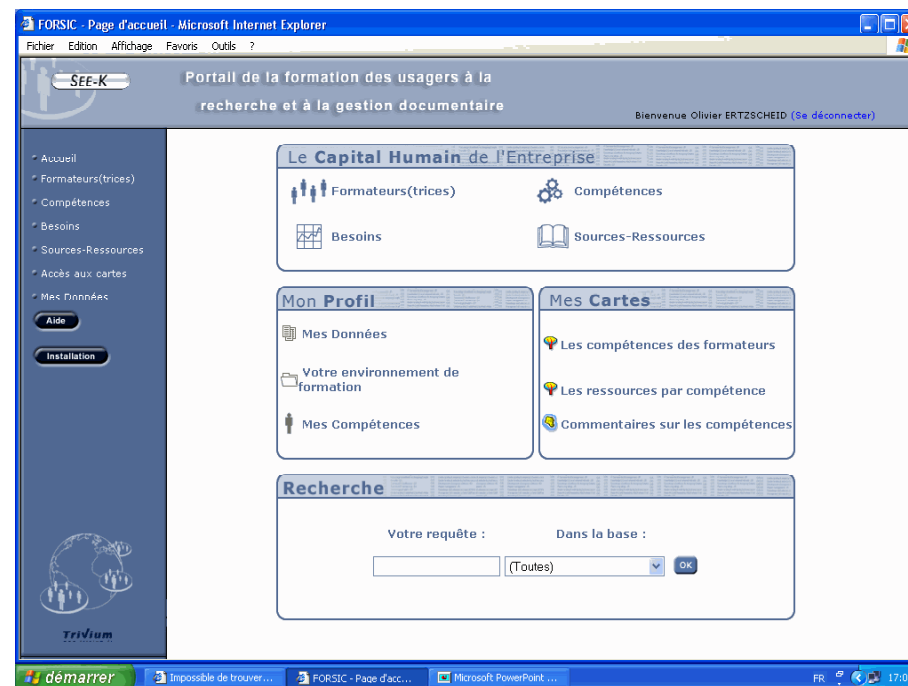


Fig. 16 : Copie d'écran de la plateforme SEE-K.

<sup>67</sup> voir le point 6.5.3. « Dimension cartographique et arbres de connaissance. »



Actuellement, comme ce travail a jusqu'ici tenté d'en faire état, le rapport de l'homme au savoir déployé dans une organisation hypertextuelle, se définit plus que jamais comme une relation de co-spécification où intention, action et mémoire humaine jouent un rôle fondamental.

Le problème des environnements actuels de coopération est que l'on y travaille avec de plus en plus de gens, sans disposer de modèles opératoires rendant compte de la globalité des échanges qui peuvent s'y produire, ni d'outils ou de solutions logicielles permettant de les traiter et de les représenter en synergie et en contexte. FoRSIC repose sur le couplage de trois solutions logicielles issues de champs scientifiques différents (multi-agent<sup>63</sup>, technologies éducatives, ingénierie des connaissances). Ces technologies sont convoquées parallèlement pour permettre :

- aux formateurs (membres du collectif de projet) :
  - d'avoir une meilleure perception de leur place dans un collectif en fonction de leurs compétences. Nous avons à cette fin mis au point un référentiel professionnel ET pédagogique des compétences en formation pour la recherche d'information ;
  - de gérer collectivement les ressources pédagogiques qu'ils ont créés ou qu'ils utilisent ;
  - de disposer d'un inventaire dynamique et évolutif de leurs sources d'information préférées ;
- aux étudiants de campus<sup>64</sup> :
  - de mieux identifier leurs manques et leurs besoins en fonction de niveaux de formation prédéfinis et validés par le collectif de formateurs ;
  - de repérer et de constituer des ressources de formation adaptées à leurs besoins individuels ;
- aux institutions et aux organismes qui choisiront d'utiliser cette plate-forme<sup>65</sup> :
  - de favoriser les collaborations entre les différents intervenants ;
  - de mettre en place ou d'ajuster une politique cohérente de formation.

## **6.2. Architecture fonctionnelle d'un dispositif de cognition distribuée.**

FoRSIC repose sur un collectif de cinquante formateurs répartis sur différents campus universitaires en midi-pyrénées.

### **6.2.1. Une plate-forme pour la gestion des connaissances.**

Ce collectif est invité, via la plate-forme web du projet constituée par le logiciel de gestion des connaissances SEE-K<sup>66</sup>, à déposer des ressources de formation, à décrire des sources de formation, à remplir

<sup>63</sup> sur la description de cette architecture multi-agent, voir [Ertzscheid et al. 01a] et [Ertzscheid et al. 01b].

<sup>64</sup> l'ouverture à cette population ne peut être faite qu'après la phase de mise au point, d'expérimentation et de validation du projet, laquelle s'achève en Décembre 2002. Ces aspects seront donc ici abordés au titre de perspectives.

<sup>65</sup> il s'agit là encore de perspectives de déploiement : un comité de pilotage (représentant les institutionnels) s'est réuni à plusieurs reprises au cours du projet mais ne s'est pas encore « approprié » le dispositif final.

<sup>66</sup> développé par la société TRIVIUM (<http://www.Trivium.fr>)

Nous voulons ici montrer en quoi l'organisation hypertextuelle est une manière originale et pérenne de questionner les bénéfices possibles de ce fantasme, par la mise en place d'une archivistique « topologique » du savoir, à la confluence du rêve borgésien de la bibliothèque universelle et de la notion foucauldienne d'archive.

A cette fin, il nous faut d'abord poser la dimension oxymorique mais non contradictoire de ces deux approches : la bibliothèque borgésienne contient tous les livres quand l'archive foucauldienne est une bibliothèque sans livres.

*« Quand on proclama que la Bibliothèque comprenait tous les livres, la première réaction fut un bonheur extravagant. Tous les hommes se sentirent maîtres d'un trésor intact et secret. Il n'y avait pas de problème personnel ou mondial dont l'éloquente solution n'existât quelque part, dans quelque hexagone. L'univers se trouvait justifié (...) mais les chercheurs ne s'avaient pas que la probabilité pour un homme de trouver sa justification, ou même quelque perfide variante de la sienne, approche de zéro. » [Borges 65 p.76]*

Pour [Foucault 94 p.499] la perspective est inversée et c'est parce qu'« *On devrait tout lire, tout étudier (...) [qu'il faut] avoir à sa disposition l'archive générale d'une époque à un moment donné. Et l'archéologie est, au sens strict, la science de cette archive.* »

La finalité utopiste de ces deux approches est aujourd'hui au cœur de problématiques bien réelles.

Concernant le projet d'une bibliothèque universelle de Borges, on sait que les projets d'élaboration d'une bibliographie universelle<sup>19</sup> furent nombreux et envisagés tout à fait sérieusement par des organismes comme l'Unesco. S'ils furent pour la plupart progressivement abandonnés, ce n'est pas tant à cause de l'ampleur de la tâche et des moyens à mettre en œuvre – problèmes auxquels des solutions matérielles auraient pu être apportées – qu'à cause d'un singulier changement de perspective dans les mentalités. En effet, ces projets ne pouvaient jusqu'alors que difficilement se départir de la consonance utopiste qu'ils véhiculaient. Mais avec l'arrivée de l'hypertexte, l'utopie devient – sous l'effet d'une alchimie mal maîtrisée – l'a-topie. L'omniprésence du livre, sa désormais évidente ubiquité est réalisée hors de la matérialité des murs de la bibliothèque, en un lieu qui n'en est pas un, un lieu qui est l'absence même de lieu puisqu'il est constitué simultanément et instantanément de tous les autres lieux possibles et existants<sup>20</sup>. Même si l'esprit demeure inapte à concevoir la notion d'infini, il ne peut qu'assister à la mise en œuvre désormais effective des réseaux infinis du savoir et de la connaissance.

Si la bibliothèque borgésienne demeure une vue fantasmée de la bibliothèque universelle, ce n'est pas tant par sa finalité que par son caractère a-topique. Depuis sa création la bibliothèque est en effet, par essence, un « *dispositif topographique* » :

*« Grâce à ce dispositif topographique de la bibliothèque d'où personne n'est exclu et où rien n'est interdit au regard analytique, se peuvent produire ces « expériences mentales » qui seront l'axe majeur de la révolution scientifique. La bibliothèque par son rassemblement ordonné multipliera la possibilité de synthèses partielles, en diffusera les compte-rendus pour élaborer une adhésion toujours rémissible et rehaussée de sa propre critique. » [Damien 95 p.111]*

<sup>19</sup> préalable indispensable à une bibliothèque universelle.

<sup>20</sup> semblable donc à l'aleph borgésien (voir chapitre premier, point 3.2.3. « In-spection »).

Et, une fois admise la disparition du Livre, la bibliothèque s'efface devant la figure de l'archive, qui, parce qu'elle se développe conjointement et en adéquation parfaite avec l'ensemble des connaissances présentes à un moment donné (et non plus à un endroit donné comme dans les rayonnages de la bibliothèque), offre ainsi à cet ensemble de connaissance un espace topologique nécessairement adapté : la carte que dessine l'archive est à l'échelle du territoire<sup>21</sup> de connaissance dont elle permet de rendre compte.

*« Par archive, j'entends d'abord la masse des choses dites dans une culture, conservées, valorisées, réutilisées, répétées et transformées. Bref toute cette masse verbale qui a été fabriquée par les hommes, investie dans leurs techniques et leurs institutions, et qui est tissée avec leur existence et leur histoire. Cette masse de choses dites, je l'envisage non pas du côté de la langue, du système linguistique qu'elles mettent en œuvre, mais du côté des opérations qui lui donnent naissance. Mon problème pourrait s'énoncer ainsi : comment se fait-il qu'à une époque donnée on puisse dire ceci et que jamais cela n'ait été dit ? C'est, en un mot, (...) l'analyse des conditions historiques qui rendent compte de ce qu'on dit ou de ce qu'on rejette, ou de ce qu'on transforme dans la masse des choses dites. »* [Foucault 94 p.786]

L'hypertexte, l'étude des modèles (« *patterns* ») sur lesquels il s'érige, le mode de distribution des agencements collectifs d'énonciation qui le parcourent, est un moyen d'analyser (en fait de spatialiser, et d'analyser les répartitions qui se donnent alors à lire) cet ensemble d'opérations donnant naissance à l'information et à son ingénierie (« *conservées, valorisées, réutilisées, répétées et transformées* »). Voilà pourquoi la notion foucauldienne d'archive – à la condition que la concordance qu'elle établit entre la carte qu'elle trace et le territoire dont elle rend compte soit vérifiée – nous semble être la clef de voûte de la mise en œuvre d'une pragmatique de la connaissance.

<sup>21</sup> sur cette distinction entre carte et territoire, voir le point 5 de ce chapitre.

- et enfin la dimension cartographique : ce projet mobilise des techniques de cartographie avancées qui reposent sur les « arbres de connaissance » développés par Michel Authier, Michel Serres et Pierre Lévy.

Nous entrerons enfin dans ce qui constitue le cœur de développement du projet, aussi bien d'un point de vue technique que théorique ou « philosophique », la construction et l'utilisation d'ontologies (point 7). Ces dernières ont souvent été évoquées dans le corps de ce travail, de manière jusqu'alors allusive car il nous semblent qu'elles nécessitent d'être présentées de manière approfondie, ce que nous nous efforcerons ici de faire.

Après avoir défini leur statut dans le cadre de la mise en place d'une organisation hypertextuelle et le rôle qu'elles sont amenées à y jouer (nous reviendrons notamment sur les derniers développements du web sémantique), nous décrirons en détail les différents niveaux ontologiques présents dans FoRSIC et les résultats concrets que permettent d'opérer certains couplages entre ces différents niveaux. Nous relaterons comment nous pensons avoir démontré que la mise en œuvre effective d'une pragmatique de la connaissance ayant pour cadre une plate-forme d'ingénierie des connaissances, nécessite certains couplages structurels entre des niveaux ontologiques déterminés, couplages ne pouvant prendre place que dans un mode d'organisation de type hypertextuel.

Il nous restera alors, avant de conclure ce travail, à remettre en perspective l'ensemble des résultats scientifiques que ce projet a permis d'obtenir, en questionnant la portabilité de ces aspects à l'ensemble de l'organisation hypertextuelle (point 7.4.4.3.).

### **6.1. Présentation du projet.**

FoRSIC est un projet de formation des usagers à la recherche d'information. Il repose sur une plate-forme de gestion des connaissances d'un collectif de formateurs œuvrant dans ce domaine.

La question de la production et de l'accès à l'information scientifique et technique dans un environnement hypertextuel est au cœur de nos préoccupations. Le projet FoRSIC choisit de l'envisager notamment, mais non exclusivement, sous l'angle de la formation des usagers. L'URFIST de Toulouse, maître d'œuvre du projet, est depuis des années fortement impliqué dans la formation de formateurs à la recherche et à l'usage de l'information : à ce titre il a mis au point un cadre théorique et méthodologique opératoire et peut compter sur un réseau de formateurs sensibilisés à ces questions et à ces méthodologies.

Dans le projet FoRSIC, nous nous sommes particulièrement intéressés aux rapports entretenus par l'information, la connaissance, les savoirs et le rapport au monde qu'ils instaurent, ainsi qu'aux rapports entre mémoires humaines (individuelles et collectives) et mémoires documentaires, et à l'évolution des pratiques et des usages en tant qu'étapes de l'externalisation de la connaissance, qui a vu la séparation progressive de la mémoire et du corps.

## 6. FoRSIC.

« L'établissement de nouvelles institutions éducatives (...) ne doit pas débiter par la question 'Que devrait-on apprendre ?' mais avec la suivante 'Avec quels types de choses et de gens les apprenants veulent-ils être en contact pour apprendre ? ' » [Illich 76 p. 80]

« La véritable valeur d'un réseau réside moins dans l'information qu'il transporte que dans la communauté qu'il forme. » [Negroponte 95 p.226]

Le projet FoRSIC<sup>58</sup> a occupé une place centrale tout au long de la rédaction de cette thèse et en constitue la partie expérimentale et applicative<sup>59</sup>. Il s'agit d'un dispositif empirique. Il a notamment permis de créer les conditions d'établissement d'une organisation hypertextuelle telle que définie jusqu'ici et de mesurer le rôle joué par cette dernière dans l'établissement et l'optimisation des conditions pouvant présider à la mise en œuvre d'une pragmatique de la connaissance<sup>60</sup>.

Après une présentation des grands axes du projet et de son architecture fonctionnelle, nous reviendrons sur ses fondements théoriques (connexionnisme et systémique), le cadre méthodologique qu'ils offrent et la manière dont nous nous y sommes inscrits pour partie ou en totalité.

Nous présenterons alors certains des principes issus des cadres théoriques précédents, comme le couplage structurel, la cognition distribuée, la « légitimation de la participation périphérique » (LPP) ... et verrons ceux qui ont pu être effectivement mis en œuvre – et ce qu'ils ont apporté au projet – et ceux non encore opérationnels mais qui restent présents et déterminants dans notre recherche d'un point de vue expérimental<sup>61</sup>.

Nous reviendrons ensuite sur les trois dimensions fondatrices de ce projet :

- la dimension coopérative : nous verrons notamment en quoi ce projet a permis de créer les conditions d'une « coopération idéale »<sup>62</sup> ;
- la dimension collective : nous verrons de quelle manière s'organisent les agencements collectifs d'énonciation entrant dans le cadre du projet ;

<sup>58</sup> FoRSIC : Formation et Recherche en Sciences de l'Information et de la Communication. Ce projet fait l'objet d'un financement du Ministère de l'éducation, de la recherche et de la technologie sous le numéro de contrat 99 – K – 6403. Site web de présentation du projet <http://www.urfist.cict.fr/forsic>. Plate-forme expérimentale du projet <http://forsic.univ-tlse1.fr>.

<sup>59</sup> la mise en œuvre de ce projet et les résultats auxquels il a permis d'aboutir et que nous décrivons dans ce sixième point sont le fruit d'un travail avec Mmes Link-Pezet et Lacombe-Caraud, responsables de l'URFIST (Unité Régionale de Formation à l'Information Scientifique et Technique) de Toulouse. Le « nous » qui sera utilisé dans ce sixième point ne désignera donc plus seulement l'auteur de ces lignes.

<sup>60</sup> l'ensemble du projet comporte des points de développement n'entrant pas dans le cadre de la problématique ici traitée (notamment le recours à des technologies multi-agent). Sur ces points nous renverrons le lecteur à des publications existantes.

<sup>61</sup> l'ensemble du projet repose sur des principes théoriques (énaction, action située et émergence notamment) qu'il ne nous est pas apparu pertinent de développer ici, pour ne pas alourdir la trame conceptuelle déjà relativement dense de ce travail. Nous les présentons sommairement dans l'annexe 12 « Principes théoriques du projet FoRSIC » et renvoyons à lecture de [Ertzscheid & Link-Pezet 01] pour une vue plus complète.

<sup>62</sup> voir le point 7.6.4. « Invariants procéduraux et déclaratifs. » du chapitre deux.

## 3. Le lieu du texte.

« Trois fois, Dieu se manifeste à Moïse dans des lieux différents. Toute parole défend son lieu. » [Jabès 91 p.117]

L'objectif que nous nous fixons maintenant est d'explorer le territoire de cette archive foucauldienne de la connaissance, en posant la question du lieu dans lequel les textes la composant s'organisent et prennent sens, tant sur leur échelle propre de signification que par rapport à l'ensemble de ceux auxquels ils sont liés ou avec lesquels ils coexistent au sein d'une entité homogène (l'organisation hypertextuelle).

Le lieu du texte est directement corrélé à la question de la construction du sens. Il est à la confluence de l'écran et de la bibliothèque sans appartenir vraiment ni à l'un ni à l'autre. Il dispose de résonances différentes si on l'aborde sous l'angle de la littérature, des aspects sociaux de la connaissance qu'il stigmatise ou des aspects plus techniques de la documentation. Nous allons nous efforcer de rendre compte de ces trois points de vue en soulignant la cohérence épistémologique que permet de leur assigner l'organisation hypertextuelle. Nombre des aspects évoqués ici ayant été développés précédemment, nous nous contenterons de rattacher explicitement ceux-ci à la problématique du lieu et renverrons aux passages concernés.

### 3.1. Littérature et espace.

De nombreux de critiques et littérateurs se sont exprimés sur les rapports particuliers qu'entretiennent la littérature et l'espace, notamment mais non exclusivement dans une optique structuraliste. Nous voulons ici reprendre et commenter l'une de ces approches qui nous paraît la plus complète, celle de [Genette 69 p.44 et suiv.]. Il distingue entre « *spatialité du langage* », « *de la lecture* », « *espace sémantique* » et « *spatialité ... temporelle* ».

#### 3.1.1. Spatialité du langage.

« Il y a tout d'abord une spatialité en quelque sorte primaire, ou élémentaire, qui est celle du langage lui-même. On a remarqué bien souvent que le langage semblait comme naturellement plus apte à 'exprimer' les relations spatiales que toute autre espèce de relation (et donc de réalité), ce qui le conduit à utiliser les premières comme symboles ou métaphores des secondes, donc à traiter de toutes choses en terme d'espace, et donc encore à spatialiser toutes choses (...) Cette spatialité se trouve en quelque sorte manifestée, mise en évidence, et d'ailleurs accentuée, dans l'œuvre littéraire, par l'emploi du texte écrit. » [Genette 69 p.44]

Cette spatialité première prend, dans le cadre de la littérature hypertextuelle, toute sa « place ». Ainsi la détermination de genres littéraires hypertextuels tient principalement à la nature des formes de l'espace narratologique qui est donné à lire, et aux contraintes liées à ces formes<sup>22</sup>.

<sup>22</sup> voir le point 8 « La forme des genres. » du chapitre premier.

### **3.1.2. Spatialité de la lecture.**

*« On peut (...) dire que l'espace du livre, comme celui de la page, n'est pas soumis passivement au temps de la lecture successive, mais qu'en tant qu'il s'y révèle et s'y accomplit pleinement, il ne cesse de l'infléchir et de le retourner, et donc en un sens de l'abolir. » [Genette 69 p.46]*

Permettant déjà de renforcer et d'inscrire au plus près de ses mécanismes de production, l'espace du langage se manifestant dans les textes, l'hypertexte confirme et institue l'abolition de la linéarité de la lecture<sup>23</sup>.

### **3.1.3. Espace sémantique.**

*« (...) l'espace sémantique qui se creuse entre le signifié apparent et le signifié réel abolissant du même coup la linéarité du discours. C'est précisément cet espace, et rien d'autre, que l'on appelle, d'un mot dont l'ambiguïté même est heureuse, une figure : la figure, c'est à la fois la forme que prend l'espace et celle que se donne le langage, et c'est le symbole même de la spatialité du langage littéraire dans son rapport au sens. » [Genette 69 p.47]*

Nous renvoyons ici à notre analyse de la rhétorique hypertextuelle, qui rattache chacune de ses figures à des espaces de signification distincts : l'ancre, le nœud cible ou pour certaines, plus rares, le lien dans son ensemble<sup>24</sup>.

### **3.1.4. Spatialité ... temporelle.**

*« Le dernier mode de spatialité que l'on peut évoquer concerne la littérature prise dans son ensemble, comme une sorte d'immense production intemporelle et anonyme. (...) La bibliothèque : voilà bien le plus clair et le plus fidèle symbole de la spatialité de la littérature. (...) On peut en dire ce que Proust dans, son 'Contre Sainte-Beuve', écrivait du château de Guermantes : 'le temps y a pris la forme de l'espace'. » [Genette 69 p.47]*

L'archivistique foucauldienne et la bibliothèque borgésienne font état de cette alchimie particulière qu'autorise et rend perceptible l'hypertexte : qu'il s'agisse de la littérature perçue comme entité homogène, de l'une ou l'autre de ses formes (genres), ou des instances d'énonciation amenées à s'y exprimer, la temporalité de type session inaugure un espace – de lecture, de navigation, de production de sens – sans cesse reproductible. *« Bolter utilise le terme d'écriture topographique : « L'écriture électronique est à la fois une description verbale et visuelle. Ce n'est pas l'écriture d'un lieu, mais plutôt l'écriture à l'aide de lieux, au moyen de thèmes spatialement réalisés. » » [Carter 97 p.17]<sup>25</sup>*

<sup>23</sup> voir notamment le point 3 « Lecteur(s) et lectures. » du chapitre premier.

<sup>24</sup> voir le chapitre second, notamment le point 5.4. « La rhétorique du lien. »

<sup>25</sup> voir aussi le point 4.1. « Les nouveaux masques de l'auteur. » du chapitre premier faisant mention de « l'information mapping » désignée par R. Horn comme l'une des rares techniques d'écriture hypertextuelle.

## Section C

- pour les secondes, seule la multiplication exponentielle des nouveaux territoires issus du processus de déterritorialisation amorcé, semble pouvoir rendre compte de la multiplicité des discours ayant effectivement cours sur les réseaux.

Dans l'un comme dans l'autre cas, le multiple répond au multiple. Or deux critères peuvent être posés comme permettant de vérifier l'adéquation entre une carte et un territoire :

- l'unique doit permettre de rendre compte du multiple dans un sens comme dans l'autre (une carte de population peut rendre compte de plusieurs territoires, un territoire peut être cartographié sous plusieurs angles) ;
- il doit y avoir une différence significative d'échelle entre la carte et le territoire, faute de quoi la carte, en devenant aussi précise que le territoire perd ses fonctions d'orientation en même temps que le territoire perd son besoin de représentation.

Cette troisième approche, en proposant des cartographies de discours prenant place au sein de dispositifs territoriaux, doit donc permettre de croiser les critères présents dans les deux premières, et confirmer qu'une organisation, quelle qu'elle puisse être, est co-déterminée par sa nature autant que par la fonction qu'elle occupe et remplit. Les points six et sept qui terminent ce chapitre sont consacrés à la description du dispositif empirique « FoRSIC » comme terrain d'expérimentation des conditions de mise en œuvre d'une pragmatique de la connaissance dans un environnement hypertextuel. Nous y décrivons la manière dont nous avons été amené à mettre en place ces croisements sur lesquels repose notre approche, c'est-à-dire des cartographies de discours non pas indépendantes mais « au service », « rendant compte », de logiques territoriales perceptibles dans les dispositifs socio-organisationnels produisant ces discours.

### **3.2. Le texte comme espace sociologique.**

Comme nous l'avons vu dans la partie consacrée aux générateurs de textes et aux formes littéraires dans lesquelles ils prennent place (L.A.O., L.G.O.), comme en a fait état le traitement des problématiques liées au versioning, c'est dans les choix d'affichage, de création et de conservation des textes numériques produits que se met en place, à l'échelle de l'organisation hypertextuelle, une sociologie des textes dont la première caractéristique est d'être non-concensuelle.

*« Cette question de la virtualité est certainement, parmi toutes les frayeurs soulevées par les arts électroniques, la plus grande car elle admet l'existence de la mort de l'objet et peut même aller jusqu'à organiser l'objet en conséquence. Or une société qui conçoit la mort des objets qu'elle engendre est une société qui admet la possibilité de sa propre disparition (...), ce qui entre dans une ligne de pensée exactement inverse à celle d'équilibre statique mise en place dans les sociétés occidentales. » [Bootz 97]*

En autorisant la dérive constante de ce socle culturel que constituait le Livre, en faisant le choix de la session, du fractal et donc de l'éphémère, l'organisation hypertextuelle des textes atteste de la remise en cause de cet « équilibre statique », linéaire, sur lequel se sont construites la plupart de nos sociétés occidentales. Ce point sera détaillé quand nous verrons comment la mise en place d'une pragmatique de la connaissance peut permettre à chacun de « penser » un nouvel équilibre plus complexe<sup>26</sup>.

*« Si on est trop jeune on ne juge pas bien, trop vieil de même. (...) Ainsi les tableaux vus de trop loin et de trop près. Et il n'y a qu'un point indivisible qui soit le véritable lieu. Les autres sont trop près, trop loin, trop haut, trop bas. La perspective l'assigne dans l'art de la peinture, mais dans la vérité et dans la morale, qui l'assignera ? » [Pascal 62 p.50]*

Laissant humblement de côté la vérité et la morale, nous pensons que pour ce qui est du domaine de la textualité, l'archivistique foucaldienne reposant sur une typologie des liens et mise en œuvre dans le cadre d'une pragmatique de la connaissance – c'est-à-dire, comme nous le verrons, d'un certain type de savoir – permettra de lui assigner son véritable lieu.

### **3.3. Le texte comme lieu technologique.**

Dans l'ensemble des technologies d'écriture et de production de textes mobilisées par l'hypertexte, c'est d'abord l'ancrage en un lieu qui est privilégié, qu'il s'agisse du lieu-adresse auquel le texte se rattache, des lieux-récit qu'il organise, ou du lieu-écran sur lequel il s'affiche. Et cette tendance ne fait que se confirmer au vu des dernières spécifications informatiques disponibles :

*« Si les premiers documents HTML comportaient avec eux leurs ancres et constituaient un document 'auto-suffisant', la multiplication des appels de scripts (Imagemap, formulaires, génération automatique de pages W3 en fonction du contexte ...) tend au contraire à localiser fortement un document. » [Le Crosnier 95]*

<sup>26</sup> voir la conclusion de ce travail.

Dernière en date, dans le cadre du web sémantique, la mise au point d'un nouveau protocole d'adressage (URI) appelé à remplacer les URL existants, et qui permettra d'attribuer à chaque ressource, à chaque texte produit, un lieu unique et déterminé une fois pour toutes<sup>27</sup>.

En plus et en parallèle de ces procédures d'adressage, l'hypertexte dispose d'une palette de composants hors-langue qu'il fait entrer dans le cœur des processus d'écriture en rendant possible un « cinétisme de l'inscription » :

*« Le 'déjà et toujours là' de l'œuvre sur le support laisse place à une surface qui est en gestation. Au lieu d'être présent, le texte arrive, surgit, apparaît. Son apparition peut-être globale et instantanée, ou bien peut se présenter sous la forme d'un enchaînement d'arrivées contrôlées de ses éléments juxtaposés ou non. Le contrôle s'exerce aussi bien sur les intervalles de temps entre l'arrivée de ces éléments que sur l'emplacement de l'inscription, ainsi que sur la ou les directions de son déploiement. »* [Papp 95].

Chacune de ces temporalités de surgissement et d'affichage est ainsi doublement localisée : par l'adressage auquel elle renvoie et par l'espace écran qui lui est dévolu.

Avec la temporalité, viennent ces deux autres constituants que sont la topographie<sup>28</sup> et le mouvement<sup>29</sup>.

*« Le texte est immuable sur le papier ; par contre, sur l'écran il a une durée, des temps forts, des temps faibles. D'abord il y a le temps de l'attente, puis le temps (le moment) de l'apparition, ensuite le temps de présence qui est la somme des temps d'états statiques, cinétiques, scintillants ou non-scintillants, latents ou réels du texte, ensuite vient le temps (l'instant) de la disparition qui est généralement suivi par un temps d'écho. »* [Papp 95]

L'étymologie de l'inscription nous renvoie à une écriture du dedans (in-scription), une écriture qui prend naissance à l'intérieur d'un objet (le texte) et d'une subjectivité (l'auteur). L'hypertexte décline cette inscription selon les mêmes modes qui lui avaient permis de décliner les modalités du regard porté sur lui par la lecture<sup>30</sup>. L'écran, surface non-inscriptible directement, n'est que la médiation, le reflet, la fenêtre ouverte sur une inscription décalée, en-deçà, rapportée, importée d'un autre lieu. Seule certitude qu'il autorise : le lieu de l'inscription hypertextuelle n'est pas l'écran. Il est ailleurs.

### 3.4. Aporie du lieu : texte utopique ?

*« La technologie électronique ôte ou soustrait du texte l'écrivain et le lecteur. (...) Il y a tellement de niveaux auxquels ils sont déferés que le lecteur ou l'écrivain éprouvent de la difficulté à identifier le lieu du texte : est-il sur l'écran, dans la mémoire de l'ordinateur ou sur le disque dur ? »* Bolter D.J. Cité par [Landow 92]

L'écran, l'archive de l'ensemble des textes disponibles, l'adresse (URI) à laquelle ils sont accédés, la mémoire (locale) de l'ordinateur sur lequel ils sont parcourus, le support physique qui les contient (cédérom, disquette), telle ou telle des versions qui les composent ... la question du lieu du texte, quel que

<sup>27</sup> ce qui rend caduque la perspective de [Lévy 88 p.17] pour qui – ce qui était vrai à l'époque – « Certes, il est possible d'assigner une adresse à un fichier informatique. Mais à l'heure de l'information en ligne, cette adresse serait de toute façon transitoire et de peu d'importance. »

<sup>28</sup> voir le point 4 « Typologie et topologie. » de ce chapitre.

<sup>29</sup> voir notamment le point 3.5. « Le mouvement de la lecture. » du chapitre premier.

<sup>30</sup> voir le point 3.2. « La lecture comme coopération. » du chapitre premier.

Organisation hypertextuelle	
Fonction	Nature
- Discours	- Territoire
- Dispositif	- Carte

Tableau 12 : Fonction et nature de l'organisation hypertextuelle.

Comme le résume ce tableau, ces deux optiques, ces deux voies de l'analyse, témoignent de postulats tacitement partagés par la plupart des réflexions engagées par rapport à cette question et auxquelles nous avons fait écho dans ce travail. Elles reprennent en l'élargissant, la distinction opérée dans notre état de l'art, entre les approches orientées-information et celles se rattachant davantage au domaine cognitif.

Pour les premières, l'hypertexte est avant tout un dispositif, qui bien que différent des bases de données et autres bases de connaissances partage cependant avec ces dernières un nombre plus ou moins grand de propriétés et/ou de finalités communes. Sa vocation première est d'architecturer pour discriminer, c'est-à-dire de faciliter la localisation, la recherche et l'accès par le plus grand nombre à des masses de données de plus en plus vastes. Ce sont donc la dimension et la nature cartographique de l'organisation hypertextuelle qui sont privilégiées : la plupart des systèmes développés dans cette approche disposent effectivement de fonctionnalités cartographiques avancées.

Pour les secondes, l'hypertexte est d'abord un discours, dans lequel le singulier se mêle plus ou moins indistinctement au collectif. Sa vocation première est d'architecturer, pour cette fois rassembler autour de propriétés homogènes et globales des discours hétérogènes et localement distribués. Pour les tenants de cette approche – parmi lesquels nombre de sociologues et de philosophes – il importe de s'interroger sur les modes d'existence et sur les modes de production communautaires de ces discours, c'est-à-dire, *in fine*, sur les territoires auxquels ils se rattachent initialement et sur ceux qu'ils contribuent à fonder en se détachant ou se démarquant des premiers : derrière la logique de déterritorialisation que la plupart de ces approches intègrent comme postulat de départ, se donne à lire la quête sans cesse renouvelée d'une reterritorialisation idéale, parce qu'en rétroaction permanente et constante avec le mouvement initial.

*« (...) le livre n'est pas image du monde, suivant une croyance enracinée. Il fait rhizome avec le monde, il y a évolution parallèle du livre et du monde, le livre assure la déterritorialisation du monde, mais le monde opère une reterritorialisation du livre, qui se déterritorialise à son tour en lui-même dans le monde (s'il en est capable et s'il le peut). »* [Deleuze & Guattari 80 p.18]

Nous plaçons en faveur d'une troisième approche, permettant d'atteindre une adéquation parfaite entre la carte et le territoire, ce que les précédentes ne permettent pas de faire :

- pour les premières, toute cartographie se met en place au vu de certains critères prédéterminés (pertinence, taux de rappel, filtres linguistiques, etc ...) qui excluent de fait de l'espace de représentation une masse importante des données par ailleurs effectivement disponibles ;

désormais du côté des singularités qui tissent son espace, qui – parce qu'elles sont le reflet d'agencements collectifs d'énonciation disposant de nouvelles organisations mémorielles, et parce qu'elles sont toutes liées à un niveau d'échelle donné – reconfigurent et retracent en permanence l'ordre du savoir, qui hérite alors de l'immanence première de ces unités.

Ainsi, pas plus qu'il ne nous paraît fondé de parler, à propos de l'organisation hypertextuelle, d'une terre de culture que d'un territoire de savoir, elle nous paraît par contre pouvoir servir de cadre (topologique) à la mise en place (à la mise en carte) d'un espace de connaissance, espace sur lequel il appartient à chacun, en son nom propre et en celui de la collectivité dans laquelle il s'inscrit, d'initier ou de rassembler les éléments épars d'une pragmatique de la connaissance. A son tour, cette pragmatique de la connaissance rend compte de l'unique territorialité de fait de l'organisation hypertextuelle.

### **5.3. Cartes fractales d'un territoire rhizomatique.**

*« (...) le rhizome se rapporte à une carte qui doit être produite, construite, toujours démontable, connectable, renversable, modifiable, à entrées et sorties multiples, avec ses lignes de fuite. » [Deleuze & Guattari 80 p.32]*

*« Nos sociétés ne fonctionnent plus à base de codes et de territorialités, mais au contraire sur fond d'un décodage et d'une déterritorialisation massive. » G. Deleuze. Cité par [Donguy 95]*

Etant donné la nature rhizomatique de l'organisation hypertextuelle et les propriétés fractales des unités qui la composent, la réponse à la question : « Quelle est la meilleure cartographie possible pour un territoire donné ? » semble être celle d'une carte fractale d'un territoire rhizomatique. Nous terminons notre argumentaire en questionnant la possibilité d'une telle carte et les conditions d'existence d'un tel territoire, ainsi que leurs implications sur la nature et la fonction de l'organisation hypertextuelle dans son ensemble.

Nous avons déjà<sup>57</sup> signalé l'importance du choix à faire entre la part du dispositif et celle du discours, ou plus exactement de la place à leur accorder. C'est dans l'équivalence qui sera établie ou dans la prédominance de l'un sur l'autre, que la fonction de l'organisation hypertextuelle des prochaines années se déterminera. C'est également en fonction de ce choix que sera déterminé le rôle dévolu aux cartes et la possibilité d'existence (et de reconnaissance) de territoires. Si l'organisation hypertextuelle fait fonction de dispositif, sa nature sera celle de la carte. Si elle fait fonction de discours, elle est alors, par nature, un territoire.

<sup>57</sup> point 8.5. « La forme des genres. » du chapitre premier.

soit l'aspect que l'on choisisse de retenir, semble constamment se dérober à l'analyse critique. Une des possibilités alors souvent choisie est de considérer le texte dans son actualité, dans son actualisation, c'est-à-dire à un moment donné de lecture, d'activation et d'état d'organisation interne des unités qui le composent (dans le cas des générateurs). Mais là encore, si nous envisageons l'un ou l'autre de ces trois moments, chacun peut être caractérisé par au moins deux des paramètres cités plus haut.

*« (...) ubiquitaire habitant du cyberspace, l'hypertexte contribue à produire ici et là des événements d'actualisation textuelle, de navigation et de lecture. Seuls ces événements sont véritablement situés. Quoiqu'il nécessite des supports physiques lourds pour subsister et s'actualiser, l'impondérable hypertexte n'a pas de lieu. » [Lévy 88 p.17]*

Le lieu de l'hypertexte est l'envers de chacun de ses endroits. Il est un lieu abîmé, il s'inscrit dans un espace<sup>31</sup> en permanence renouvelé de significations. C'est cet espace qui est adressable, qui peut être cartographié. C'est dans ces cartographies, dans cette approche topologique de l'espace que nous allons maintenant aborder, qu'apparaissent puis disparaissent des foyers d'émergence de la textualité, dans le temps de la perception cognitive d'une individualité de lecture ou d'une collectivité d'écriture.

La littérature fut d'abord circulaire : sa transmission orale conditionnant une certaine perception du temps. Transmission par l'exemplarité du récit.

La littérature fut ensuite linéaire. N'ayant plus à se soucier de son mode de transmission et de diffusion, assuré par la simple reproduction (imprimerie), elle se déclina sur le mode du prolongement.

La littérature est aujourd'hui cartographique. Ecriture de l'espace, dans l'espace, d'un espace.

<sup>31</sup> « (...) l'espace c'est ce qui empêche que tout soit à la même place. » [Virilio 90 p.127]

**Citations originales.****Point 1. L'écran : le dialogue du corps et du texte.**

- [Mathews 93] « *It's our architecture today : great screens on which are reflected atoms, particles, molecules in motion. Not a public scene or true public space but gigantic spaces of circulation, ventilation and ephemeral connexions.* »

**Point 2. Bibliothèque.**

- [Bush 45] « *human mind work by associations.* »

**Point 3. Le lieu du texte.**

- [Carter 97 p.17] « *Bolter uses the term topographic writing : « Electronic writing is both a visual and a verbal description. It is not the writing of a place, but rather a writing with places, with spatially realized topics.* ».
- [Landow 92] « *Electronic technology removes or abstracts the writer and reader from the text. (...) There are so many levels of deferral that the reader or writer is hard put to identify the text at all : is it on the screen, in the transistor memory or on the disk ?* »

sensible dans le cas des générateurs de texte qui, sur la base d'un mécanisme de génération unique ou disposant de contraintes aléatoires, sont tour à tour capables de générer des fragments prenant place au sein d'un même territoire (celui de l'œuvre) ou donnant lieu à une multiplicité d'œuvres et de parcours.

Le deuxième axiome (un territoire se définit par la simple collection physique de ses fragments), devient immédiatement aussi inexact qu'inapproprié, l'hypertexte étant une collection physique de fragments organisés de manière fractale.

Seule la dernière règle demeure pérenne, à cette différence près que la propriété dont il est alors question n'est plus exclusivement une propriété mesurable en terme d'appartenance ou d'inclusion (telle ville appartient à telle région qui appartient à tel pays, etc ...) mais également en termes de fonction, de nature, et de mode d'organisation, ces critères étant par ailleurs non-exclusifs et pouvant être, dans leur ensemble ou pour partie, co-dépendants.

**5.2.2. Le savoir entre terre de culture et territoire de connaissance.**

La notion de territoire est aussi fédérative que fédératrice, et il importe donc que tout objet ou tout artefact puisse être analysé sous l'angle (et au moyen des cartes) d'un territoire d'appartenance. La notion de déterritorialité, sur laquelle nous reviendrons sous peu, ne faisant elle-même sens qu'en fonction d'un nombre donné de territoires dont il est alors possible de se démarquer. Ces territoires apodictiques sur lesquels s'opère la série de renégociations dont nous venons de parler sont ceux de la « culture » et du « savoir ».

Pour [Debray & Finkelkraut 94 p.242] :

« (...) la culture est toujours territoriale : elle est toujours liée à une démarcation, à une circonscription, à une appartenance. Il n'y a de culture que locale, il n'y a de génie que des lieux, il n'y a de littérature qu'en langue naturelle et, par définition, aucune langue naturelle n'est cosmique ni mondiale (...): elle doit garder quelque chose qui n'est livré qu'à une interprétation, et non à une équivalence. » R. Debray.

S'il nous paraît effectivement fondé de parler d'un territoire d'appréhension et d'appropriation de l'hypertexte qui soit de l'ordre du « culturel », celui-ci ne nous paraît plus devoir fonctionner sur la constante culturelle pointée par Debray, constante qui, dans la pratique des discours communautaires qui se donnent à lire dans l'organisation hypertextuelle<sup>55</sup>, devient une variable soumise aux règles d'usage en vigueur dans chacune des ces communautés.

Le second territoire de référence est celui du savoir. Se pose alors la question de déterminer « *Qu'est-ce qu'un Territoire dans l'ordre du savoir ?* » et de répondre « (...) *L'imposition d'un ordre transcendant, céleste, total, sur un espace immanent et lacunaire des singularités.* » [Lévy 81 p.181]<sup>56</sup>. Or l'hypertexte, du fait de ses modes d'organisation, inverse également cette perspective : la transcendance est

<sup>55</sup> il ne s'agit en aucun cas de faire référence à un quelconque communautarisme sociologique. Par « discours communautaires », nous faisons ici référence aux conjonctions naissant de la redéfinition de la carte énonciative (agencements collectifs d'énonciation) et des nouvelles organisation mémorielles s'y rapportant (voir le point 4 « Emergence de nouvelles subjectivités. » du chapitre premier).

<sup>56</sup> voir aussi la préface de Michel Serres dans [Authier 98b].



L'utilisation et la place des techniques de cartographie dans le cadre de l'étude de l'hypertexte peut se décliner sous trois angles principaux :

- historiquement et étymologiquement, Hypercard est le premier système permettant la création d'hypertextes et a, par extension, inscrit dans l'usage courant le rapprochement entre ces deux notions<sup>52</sup> ;
- pratiquement, les techniques courantes de cartographie ont prouvé leur efficacité en même temps que leur nécessité pour pouvoir rendre compte des infrastructures matérielles et techniques de l'organisation hypertextuelle – dont peuvent être inférées nombre d'analyses sociologiques. Le site fondateur de Martin Dodge, CyberGeography<sup>53</sup>, est exemplaire à cet égard ;
- épistémologiquement enfin, l'ensemble des techniques de « mapping » faisant aujourd'hui fonction d'interface principale dans la plupart des systèmes présents sur le web (moteurs de recherche, outils de gestion de l'information et par extension de gestion des connaissances, et logiciels d'écriture hypertextuelle – StorySpace) témoigne de l'adéquation déjà soulignée entre carte et hypertexte pour l'analyse de l'ensemble des modalités individuelles et collectives de constitution des savoirs et de leur accès<sup>54</sup>.

## **5.2. Territoire(s) et territorialité(s).**

### **5.2.1. Logiques territoriales fractales.**

Il existe une logique mathématique de l'identité territoriale, proposée par J.M. Laborde et reprise par [Parrochia 91 p.196] :

- *« la notion d'identité territoriale s'exprime ainsi : A a même territoire que B si A est fragment de B et B est fragment de A ;*
- *on définit le territoire de A par la collection physique de ses fragments ;*
- *le territoire peut être conçu comme la réunion d'un certain nombre de fragments possédant en commun une certaine propriété. »*

Le premier axiome de cette logique devient caduque dans le cadre de l'organisation hypertextuelle.

Du fait de sa nature fractale, il devient possible pour une lexie « A » d'être à la fois le territoire d'une lexie « B » qu'elle contiendrait et l'un de ses fragments, reprenant alors à un niveau d'échelle différent l'ensemble des caractéristiques de « B » (la plupart des ancres hypertextuelles telles que nous les avons décrites présentent cette double caractérisation). Ce nouvel ordre logique territorial est particulièrement

<sup>52</sup> « Dans le milieu informé on parle d'abord [à propos de la page-écran] de stacks (...) puis de « cartes » ou « d'hypercartes », reprenant par là le nom d'un logiciel d'écriture hypertexte pour Macintosh (Hypercard). » [Vandendorpe 99 p.196]

<sup>53</sup> <http://www.cybergeography.org>.

<sup>54</sup> « D'où l'importance du « mapping » aujourd'hui, au cœur, non seulement des analyses concernant les réseaux socio-techniques, les mémoires numériques collectives en acte, mais encore des pratiques coopératives. Cartographe, selon diverses métriques et niveaux d'échelles les dynamiques d'actants, à partir d'un nombre croissant de traces, dont les traces « processuelles » que sont les liens et hyperliens, devient primordial, afin d'améliorer l'activité coopérative des agents humains/non-humains. La mise en jeu socio-cognitive de ces « pratiques cartographiques », au sein des intelligences collectives émergentes, posant certains problèmes, dont un des plus délicats est la renégociation politique du partage des savoirs transversaux, des légitimités institutionnelles acquises au cours de la longue histoire techno-politique, passionnelle, des savoirs protégés, des opacités cognitives et informationnelles héritées, de la construction et du partage des secrets. » [Link-Pezet & Noyer 99]

## Section B

#### 4. Typologie et topologie.

« C'est dans les époques de dislocation historique, quand nous sommes chassés du lieu, que se développe pour elle-même cette passion structuraliste qui est à la fois une sorte de rage expérimentale et un schématisme proliférant. » [Derrida 67 p.14]

« La science - autrement dit l'agrégat des messages à propos du système que l'on est en train de décrire - sera construite de telle façon qu'il sera possible de la représenter dans un diagramme plus ou moins complexe de types logiques. Telle que je l'imagine, chaque message aura son emplacement sur cette carte et la relation topologique entre différents emplacements représentera le relation typologique entre les messages. Il est dans la nature de la communication, telle que nous la connaissons, d'admettre la possibilité d'une telle carte. » [Bateson 77 p.180]

Les diverses typologies proposées dans ce travail (typologie des liens, des instances d'énonciation, des « genres » hypertextuels, des stratégies de navigation, de la littérature informatique, des réseaux ...) nous ont permis d'isoler un certain nombre de paramètres et d'invariants. Nous avons choisi, pour chacune de ces typologies, de les représenter sous la forme d'une vue synoptique et avons commencé, tout au long de notre exposé, à distinguer et à commenter nombre de corrélations et de transversalités existant entre ces différentes typologies, dont le point commun est de permettre de rendre compte des structures, des principes et des modes d'organisation et de déploiement fondateurs de l'organisation hypertextuelle. Nous avons ainsi construit, une représentation que nous pensons globale et cohérente puisqu'au vu des premières transversalités dégagées, chacune d'entre elles résiste isolément à l'analyse et confirme la vue d'ensemble qu'elles permettent d'offrir. La superposition de ces différentes typologies établit une topologie de l'organisation hypertextuelle, isolant aussi bien des mécanismes, des processus de répartition, que des lieux, des agrégats dans lesquels se condensent, par nature ou par fonction, certaines des unités ainsi réparties.

Dans le même temps, pour nombre de ces typologies (notamment celle des liens et des genres hypertextuels) notre analyse a permis de faire ressortir comme particulièrement pertinents et discriminants, un certain nombre de critères fondés sur des concepts directement issus de cette branche des mathématiques qu'est la topologie (« patterns » hypertextuels, propriété de connexité des ancrs hypertextuelles ...). Nous avons déjà commencé d'esquisser l'importance de ce champ d'ailleurs explicitement posé par Lévy comme l'un des six principes de l'hypertexte.

Nous voulons maintenant terminer le travail ainsi entrepris. Après avoir rappelé certains des principes fondateurs de la topologie, nous montrerons les liens étroits qui unissent nombre de typologies produites dans le champ des sciences de l'information et de la communication ainsi que dans celui des études littéraires avec les concepts empruntés à la topologie. Nous reviendrons enfin sur les principes topologiques caractéristiques de l'organisation hypertextuelle. La topologie étant définie par [Steiner 98 p.573] comme la science permettant d'établir « des rapports d'invariance à l'intérieur d'une transformation », nous verrons en quoi elle se prête « naturellement » à la remise en perspective des invariants jusqu'ici dégagés.

de communication n'est pas le même que celui d'une œuvre littéraire hypertextuelle, qui n'est pas le même que celui du site hébergeant cette œuvre parmi d'autres, etc.)

Après avoir examiné chaque concept – carte et territoire – indépendamment, pour en tracer rapidement la genèse et les problématiques spécifiques dans le cadre de l'hypertexte, nous établirons une série de convergences qui permettent de répondre à la question de « la meilleure carte possible pour un territoire donnée » ou à tout le moins de comprendre en quoi cette question est centrale pour l'ensemble des points que nous avons jusqu'ici traités.

#### 5.1. Carte(s) et cartographie(s).

« Sur une pancarte, au mur, le profil d'un bœuf apparaît comme une carte géographique parcourue par des frontières qui délimitent les aires d'intérêt comestible (...). C'est là une carte de l'habitat humain, non moins que le planisphère du globe terrestre : aussi bien l'une que l'autre sont en somme des protocoles sanctionnant les droits que l'homme s'est attribués, de possession, de partage et de dévoration sans résidu, des continents terrestres comme des lombes du corps animal. » [Calvino 85 p.78]

Comme cela ressort de la citation de Calvino, la carte est d'abord un « protocole », un mode d'appropriation, par la représentation, du territoire réel (et parfois fantasmé dans le cas de l'hypertexte). Si elle demeure à un niveau de signification qui a d'abord à voir avec la nature de l'objet-carte plus qu'avec celle des objets ou des entités cartographiées<sup>49</sup>, ce niveau de signification est de même nature que celui que nous avons choisi d'affecter à l'étude de l'organisation hypertextuelle : la perspective que renverse d'emblée l'établissement des premières cartes est celle de notre rapport au monde tel qu'il peut être corrélé ou inféré de notre rapport au savoir<sup>50</sup>. « Mercator, lorsqu'il publie pour la première fois son Atlas, peint dans le frontispice non plus un géant portant le monde sur son dos, mais un savant tenant la Terre dans sa main. » [Latour 89 p.364].

C'est effectivement dans cette optique que [Lévy 91 p.82] pose la carte comme une « technologie intellectuelle » :

« (...) dès le moment où, indépendamment de sa présence concrète, son image mentale est utilisée par un individu pour évaluer la distance entre deux points d'un territoire ou pour établir une stratégie quelconque. Même quand elle n'est physiquement plus là, la carte est devenue un élément essentiel de l'outillage mental d'un sujet cognitif. »

A l'image des « cartes muettes » ne portant pas de nom<sup>51</sup>, toute carte rendant compte d'un aspect donné de l'organisation hypertextuelle propose, à part égale, une échelle de significations préexistantes (celles qui fondent l'existence de l'objet représenté) combinées à d'autres, d'abord en attente et ne devenant actives que dans le cadre d'un parcours de navigation.

<sup>49</sup> la même carte peut servir à désigner des zones habitables ou non-habitable, etc. Dans l'exemple de Calvino, c'est le potentiel cartographique de l'entité bœuf qui est d'abord représenté. Les aspects comestibles ou non-comestibles des unités présentes sur la carte ne dépendent que du mode de déchiffrement qui pourra être inféré par chaque lecture, par chaque processus individuel d'orientation.

<sup>50</sup> voir le point 6.5.3. « Dimension cartographique. » de ce chapitre et la conclusion de notre travail.

<sup>51</sup> ces cartes muettes sont fréquemment utilisées dans l'enseignement, les élèves devant alors trouver et placer ces noms sur la carte.

## 5. La carte et le territoire.

« Le langage entretient avec les objets qu'il désigne le même rapport que la carte entretient avec le territoire. »  
[Bateson 77 p.212]

Le rapport problématique, posé par Bateson, que la carte entretient avec le territoire dont elle permet de rendre compte et sur lequel elle permet de s'orienter, est celui de l'existence de possibles adéquations entre un langage (courant, iconique, visuel, symbolique) et son objet. Ce rapport est à l'origine de bien des controverses en ce qu'il mêle plusieurs niveaux de réalité (concrète et physique pour le territoire, mentale et construite selon un code donnée à une échelle donnée pour la carte).

« Sans doute est-ce à partir de la découverte d'un espace composé de points abstraits que Zénon d'Elée put formuler son fameux paradoxe sur l'impossibilité du mouvement. (...) Sur le Territoire, Achille ne rattrape jamais la tortue. On n'avance pas sur le Territoire, on ne peut qu'y faire le point. Mais sur la Terre, constituée par les randonnées, les parcours et les voyages, Achille rattrape la tortue. Or Achille (...) vit dans les deux espaces. Il se repère et s'oriente sur le Territoire, mais il progresse sur la Terre. » [Lévy 81 p.180]

Probablement plus que tout autre phénomène se prêtant à une représentation cartographique, l'organisation hypertextuelle pose, à chacun de ses niveaux de réalité physique et de représentation mentale, le problème de déterminer quelle est la meilleure cartographie possible<sup>48</sup> pour un territoire donné, qu'il s'agisse du territoire qui peut être parcouru dans une session de navigation ou de celui constitué par l'ensemble de ces unités (et donc toujours à un moment donné). La carte de l'organisation hypertextuelle – et c'est là la première propriété qui la distingue de toute autre – est donc liée non seulement à un espace, à un territoire, mais également à une temporalité particulière.

La cartographie « traditionnelle » est expansive, cumulative et centrée. Expansive parce que les unités qui la composent migrent le plus souvent vers des unités plus grandes (routes devenant autoroutes, etc.) ; cumulative parce qu'elle procède le plus souvent par ajout et non par soustraction (de nouvelles voies apparaissent) ; centrée parce qu'elle a le plus souvent pour objet une réalité physique donnée du territoire (carte fluviale, cartes de densité de population, carte des types d'agriculture, etc.), ces réalités pouvant également être cumulées sans que soit remise en cause la pertinence de chacune. Ainsi, chaque objet relève du même territoire. Chaque territoire se prête à toutes les cartographies possibles.

La cartographie hypertextuelle est co-dépendante, associative et a-centrée. Co-dépendante puisque les unités liées ne le sont pas nécessairement en vertu d'un rapport de grandeur permettant des emboîtements linéaires. Associative parce qu'en plus de rendre compte d'ajouts et de soustraction d'unités, ces ajouts et ces soustractions ne font sens qu'au travers des relations associatives préexistantes à ces deux modes (la disparition/l'ajout d'une unité n'entraîne pas nécessairement la reconfiguration du micro-ensemble auquel elle appartenait mais implique par contre celle du macro-ensemble). A-centrée parce qu'à certaines des réalités physiques de l'organisation (nature des unités liées, type d'hypertexte dans lequel elles prennent place) correspondent des logiques territoriales différentes (le territoire des connexions physiques des réseaux

<sup>48</sup> ses propriétés topologiques répondant par l'affirmative à la question de savoir si une cartographie est possible.

## 4.1. Topologie : définitions.

N'étant pas mathématicien de formation, et notre lecteur n'étant pas supposé l'être, nous commençons par reprendre quelques unes des définitions qui nous sont apparues les plus éclairantes en les commentant à l'aune de notre sujet pour commencer à établir des points de correspondance sur lesquels nous reviendrons par la suite.

La consultation du dictionnaire encyclopédique Quillet nous apprend que la topologie désigne en mathématique, la « *structure destinée à formaliser et à généraliser les notions intuitives de borné, d'ouvert, de continu, de frontière, pour un ensemble de points. - Branche des mathématiques consistant à étudier les ensembles munis d'une telle structure.* »<sup>32</sup>.

La plupart des questions qui fondent la problématique de l'organisation hypertextuelle reposent elles aussi au premier plan sur ces mêmes notions intuitives, comme l'a montré notre insistance tout au long de ce travail sur l'importance et la difficulté que représente la variation constante du niveau d'échelle choisi comme angle d'analyse. Les points topologiques sont ici les différentes lexies composant l'hypertexte. Notre étude des liens hypertextes n'a poursuivi d'autre but que celui d'étudier des ensembles de textes du point de vue des structures et des degrés de fermeture et/ou d'ouverture dont ils disposent ou qu'ils permettent de mettre en œuvre.

La définition proposée par [Gleick 91 p.69] dans une optique de vulgarisation, pose une nouvelle fois la spécificité de l'organisation hypertextuelle comme relevant de la topologie. « *La topologie est la géométrie des surfaces élastiques. Elle s'intéresse davantage au qualitatif qu'au quantitatif, et demande : « Si vous ignorez les mesures, que pouvez-vous dire sur la structure globale ? » »*

Comme le confirme tout un ensemble d'études déjà évoquées, si l'hypertexte n'est pas à proprement parler une surface élastique, il n'en demeure pas moins que la topologie de cet espace est constamment renouvelée, modifiée et que les études s'efforçant d'établir des mesures quantitatives (diamètre du web) apparaissent nécessairement datées quand celles qui s'intéressent à la forme d'ensemble présentent des caractéristiques pérennes (le web en forme de nœud papillon)<sup>33</sup>.

La définition de [Tinland 91 p.195], plus applicative, confirme ce point de vue : « (...) *la topologie, étude mathématique des distorsions, plus exactement étude des propriétés qualitatives et des positions relatives des êtres de raison géométriques indépendamment des détails de leur forme et de leur grandeur.* »

Cette première série de définitions permet d'éclairer l'affirmation que nous posons dans le point 8.5. de notre premier chapitre<sup>34</sup> selon laquelle la topologie est bien une science étrange pour laquelle « (...) *une tasse à café est identique à une chambre à air, car toutes deux sont des surfaces avec un trou.* » Elle dispose en effet d'un principe selon lequel « *Pour déterminer une équivalence topologique, vous avez le droit de*

<sup>32</sup> la topologie se subdivise en trois branches : la topologie générale (qui concerne les ensembles ouverts et leurs voisinages), la topologie métrique (à laquelle nous avons déjà fait référence sans la citer au travers du principe de connexité : « *on dit qu'un espace topologique est connexe si le seul ensemble de cet espace à la fois ouvert et fermé est l'espace tout entier.* », en enfin la topologie algébrique (qui s'intéresse aux chemins – courbes – parcourant un espace topologique donné).

<sup>33</sup> voir le point 4.5.1.1. « Approches orientées information. » du chapitre deux.

<sup>34</sup> « La forme des genres : pour une critique topologique. »

*couper une forme, sous réserve de recoller les deux lèvres exactement comme elles l'étaient initialement (en effet, une même forme topologique peut-être plongée dans l'espace de différentes manières).* » [Stewart 00 p.107]

Faute de temps, de place (et de compétence) nous n'indiquons ici qu'au titre de piste de recherche, l'intérêt particulier que certains principes topologiques nous semblent offrir pour l'étude de l'organisation hypertextuelle. L'un des plus pertinents est celui dit de « *l'équivalence topologique de deux objets* » (deux lexies) :

*« Deux objets sont topologiquement équivalents (ou appartiennent à la même classe) lorsqu'il existe une bijection entre les points de la surface des deux objets : à chaque point de l'un correspond un point de l'autre. Cette définition est dite « intrinsèque », elle devient « plongée » quand une telle bijection existe aussi pour les régions délimitées – l'intérieur et l'extérieur – par les ensembles de points constituant les objets. »* [Douady & Mangin 00 p.108]

Nous aurons l'occasion de revenir un peu plus avant sur ce principe d'équivalence (qui fait référence à une appartenance de classe) dans la partie consacrée aux ontologies.

#### **4.2. La topologie comme (inter)discipline.**

##### **4.2.1. Topologie et SIC.**

Nous n'entrerons pas ici dans une étude détaillée des rapports qui existent entre topologie et sciences de l'information et de la communication. Nombre des études jusqu'ici mentionnées comme se rattachant au champ des S.I.C. reposent sur l'établissement de formes topologiques permettant d'analyser (pour optimiser des processus de recherche) la nature, la quantité d'information et le degré de liaison de l'information disponible sur les réseaux. La plupart des techniques de représentation cartographique utilisées dans divers outils logiciels de DataMining<sup>35</sup>, d'ingénierie documentaire, d'analyse lexicale ou de gestion des connaissances se fondent également sur une approche *in fine* topologique<sup>36</sup>.

La pertinence du rapprochement de ces deux champs culmine dans la proposition de [Lévy 94b p.116] plaçant pour la mise en œuvre d'une « *Topologie récursive : nouvelle discipline mathématique dont l'objet est la construction d'espaces topologiques à partir d'une configuration informationnelle.* »<sup>37</sup> qui est *de facto* l'angle d'approche choisi par certains moteurs de recherche et certains logiciels de gestion d'information<sup>38</sup>.

##### **4.2.2. Topologie et littérature.**

S'ils peuvent *a priori* paraître moins explicites, les rapports entre topologie et littérature n'en sont pas moins féconds et pertinents dans notre cadre d'étude. L'entrée de la topologie en littérature procède par

<sup>35</sup> le DataMining désigne une catégorie d'outils permettant la fouille, l'extraction de données.

<sup>36</sup> nous y reviendrons au travers de l'exemple des cartographies produites par les arbres de connaissance utilisés dans le cadre du projet ForSIC auquel nous consacrons le point 6 de ce chapitre.

<sup>37</sup> ce principe de topologie récursive a été inventé et mis au point par Michel Authier, président de la société Trivium. C'est sur ce principe que reposent les Arbres de connaissance (voir le point 6.5.3. de ce chapitre).

<sup>38</sup> <http://search.mapsatn.net>, <http://www.kartoo.com>, <http://miner3d.com>, <http://www.trivium.com> ...

Au-delà des caractérisations structurelles qui s'appliquent individuellement à chacune des unités présentes sur le réseau et qui relèvent – de manière apparemment distincte – de structures logiques, sémantiques, spatiales et temporelles, la dimension instaurée par les ancrs hypertextuelles les condense et les rassemble en même temps qu'elle autorise et organise leur déploiement à l'échelle individuelle. L'organisation hypertextuelle n'a qu'une seule dimension qui se décline en plusieurs espaces. Cette dimension, invariante, est celle qui relève des structures de liaison entre unités, qui, pour reprendre un concept également présent dans la théorie des catastrophes, fonctionnent comme autant « d'attracteurs »<sup>45</sup> opérant au niveau des styles cognitifs mobilisés lors de chaque session de navigation<sup>46</sup>.

Le principal effet de ces attracteurs est de permettre « *par un petit changement de paramètres, [d'] obtenir une infinité de types typologiques du système correspondant. Autrement dit, il y a une instabilité topologique du système en presque tous les points de l'espace de contrôle.* » [Thom 93 p.50] Cette instabilité topologique ne remet par ailleurs en cause ni les stabilités topographiques locales, ni l'existence de propriétés topologiques globales. Elle détermine, caractérise et explique la nature du phénomène de variation du niveau d'échelle.

A chaque action sur l'un de ces liens ou l'une de ces ancrs qui constituent ces « *points de l'espace de contrôle* » du système, le territoire se modifie et il faut se mettre en quête de nouvelles cartes permettant d'en dresser les limites et les points de fuite. « [les topographies] *sont des expressions performatives qui projettent et créent simultanément un territoire. Avec l'Internet, cette fonction performative est plus marquée encore puisqu'aucun « sol » rassurant ne se trouve sous l'écriture du lieu.* » [Abendroth & Decock 00b]<sup>47</sup>

<sup>45</sup> structure qui apparaît à un moment donné de l'itération d'une fonction non-linéaire.

<sup>46</sup> [Abendroth & Decock 00b] parlent de « *paysages mentaux* » ou « *brainscapes* » : « *Les environnements générés par des données organisent l'information en temps réel en reliant les flux de données, rendant perceptible ces liens à travers un espace à la fois cognitif et mental – brainscapes.* »

<sup>47</sup> il s'agit d'une citation traduite de l'article de Nunes M., « **Virtual Topographies : Smooth and Striated Cyberspace** » [en ligne] <http://www.dc.peachnet.edu/~mnunes/vtop.htm>, consulté le 10/09/2001.

#### 4.3.2.1. Topographies de navigation.

« (...) l'espace sédentaire est strié, par des murs, des clôtures et des chemins entre les clôtures, tandis que l'espace nomade est lisse, seulement marqué par des « traits » qui s'effacent et se déplacent avec le trajet. » [Deleuze & Guattari 80 p.472]

Pour [Abendroth & Decock 00b] :

*« Il existe au moins deux formes de topographies possibles, l'une linéaire et orientée-point, l'autre fluide et orientée-plan. (...) Dans les topographies striées, chaque point est fixé dans une position donnée, déterminée par des lignes et des trajectoires, produisant une organisation statique, mais efficace, tout en réduisant les interrelations spatiales en une simple communication à deux voies. De cette manière, la disposition spatiale striée a tendance à être un espace hiérarchique et arborescent, organisant l'information à travers des niveaux et des degrés. A l'opposé, le modèle de l'espace lisse est basé sur les multiples interconnexion (...), où chaque point devient un relais et existe seulement comme relais. Cette interconnexion multi-directionnelle produit un environnement plan et fluide « écrivant » le cyberspace comme environnements de transit et de passage continus – un espace nomade et mental de communication digitale. »*

Tout l'argumentaire développé ici, reprend et confirme la distinction que nous avons opérée entre les deux modalités de navigation que sont le « Browsing » (« linéaire, orienté-point, statique mais efficace, espace hiérarchique et arborescent ...») et le « Searching » (« espace lisse, chaque point devient un relais, environnement plan et fluide ...»). Ainsi, et de manière générale, l'activité de navigation dans le cadre d'une session, génère sa propre organisation hypertextuelle (topographie), de la même manière que l'organisation hypertextuelle globale met en œuvre et autorise des systèmes distincts de navigation (topologie).

#### 4.3.2.2. Les liens comme résistances topographiques.

Dans le temps de la session, les caractéristiques topographiques des unités parcourues apparaissent comme :

*« (...) des espaces homogènes, localement homogènes. Ces espaces sont ce que nous appelons des variétés. (...) Mais des singularités apparaissent lorsque l'on soumet en quelque sorte l'espace à une contrainte. La manche de ma veste, si je la comprime, je fais apparaître des plis. (...) ; lorsqu'un espace est soumis à une contrainte, c'est-à-dire lorsqu'on le projette sur quelque chose de plus petit que sa propre dimension, il accepte la contrainte sauf en un certain nombre de points où il concentre toute son individualité première. Et c'est dans la présence de ces singularités que se fait la résistance. Le concept de singularité, c'est le moyen de subsumer en un point toute une structure globale. » [Thom 93 p.23]*

L'approche de Thom développée dans le cadre de la théorie des catastrophes nous permet ici de confirmer plusieurs points déjà démontrés ou posés comme postulats dans ce travail :

- la nature fractale de l'organisation hypertextuelle est une nouvelle fois démontrée par l'existence de singularités capables « de subsumer en un point toute une structure globale »,
- ensuite, comme nous en avons fait l'hypothèse, ce sont bien les liens – au travers de leurs ancrs – qui constituent ces résistances dans lesquelles des singularités se donnent à lire et se mettent en place.

contamination, opérant d'abord au niveau le plus fin – celui de la rhétorique – pour s'étendre ensuite aux lexies, et enfin jusqu'à la notion d'œuvre.

#### 4.2.2.1. De la rhétorique à l'Oulipisme.

Il faut remonter à la rhétorique classique et à ses tropes pour trouver la formulation d'un lien explicite entre topologie et littérature. Le dictionnaire Quillet mentionne ainsi pour l'entrée « Topologie » une rubrique « Rhétorique » dans laquelle on peut lire qu'il s'agit de la « *Connaissance des lieux communs, des sources où peut puiser un prédicateur.* » De la même manière, la topographie permet, en rhétorique classique, de désigner « *la description d'un lieu par la réunion des détails qui frappent le mieux l'imagination.* »<sup>39</sup>

Il faudra attendre la constitution de l'OuLiPo<sup>40</sup> pour que la topologie réapparaisse de manière significative et de plein droit dans le champ des études littéraires. Pour François Le Lionnais, l'un de ses fondateurs :

*« Les mathématiques – plus particulièrement les structures abstraites des mathématiques contemporaines – nous proposent mille directions d'explorations, tant à partir de l'Algèbre (recours à de nouvelles lois de composition), que de la Topologie (considérations de voisinage, d'ouverture ou de fermeture des textes). » [Oulipo 73 p.17]*

Il s'agit pour l'OuLiPo de s'intéresser à « (...) la transposition dans le domaine des mots de concepts existants dans les différentes branches des mathématiques : Géométrie (poèmes tangents entre eux de Le Lionnais), Algèbre de Boole (intersection de deux romans de J.Duchateau), Algèbre matricielle (multiplications de textes de R. Queneau). » [Oulipo 73 p.46]

#### 4.2.2.2. Pour une lexie topologique.

Nous avons choisi dans ce travail de retenir la notion de « lexie », empruntée à Barthes, comme unité de référence dans le cadre de l'organisation hypertextuelle. Nous avons également souligné que d'autres unités de référence étaient possibles et avaient été proposées. La principale alternative à la lexie barthesienne est la dichotomie entre « texton » et « scripton » d'Espen J. Aarseth :

*« C'est en songeant aux propriétés du texte désormais affranchi des limites de son support que Espen J. Aarseth a pu proposer une topologie textuelle qui cherche à définir les unités du texte non plus dans leurs rapports linguistiques, mais dans leurs rapports au tout et à ses parties. (...) Il suggère d'appeler Texton l'unité textuelle de base définie dans une perspective topologique, et Scripton une séquence non-interrompue d'un ou de plusieurs textons tels qu'ils sont projetés par le texte ou réunis par le lecteur. » [Clément 95]*

<sup>39</sup> voir aussi le point 6.2. « L'image haut-lieu du texte. » de notre premier chapitre où nous évoquons les travaux de [Yates 75] à propos des « ars memoriam ».

<sup>40</sup> OuLiPo : Ouvroir de Littérature Potentielle. Pour une présentation générale de ce groupe et de ces travaux on consultera [Oulipo 73]. L'OuLiPo occupe une place importante dans la mise en place de nombre de pistes théoriques et d'œuvres littéraires préfigurant ce qu'est actuellement la réalité de l'organisation hypertextuelle (que l'on songe aux **Cent mille milliards de poèmes** de Queneau, à certaines de œuvres de Pérec ou de Roubaud). Certaines de ces préfigurations apparaissent dans notre typologie de la littérature informatique. Nous exprimons ici notre second regret qui est de n'avoir pas eu le temps d'amener le lecteur en « plongée » hypertextuelle dans certaines des œuvres oulipiennes de référence. Regret tempéré par l'abondance et la qualité des ouvrages et études par ailleurs consacrés à l'analyse de ces œuvres.

D'unité de lecture qu'elle était chez [Barthes 70 p. 18], la distinction entre texton et scripton fait davantage référence à un parcours de navigation permettant de faire sens en amont du processus de lecture. Elle confirme l'ancrage topologique déjà perceptible chez Barthes, parlant « *d'un meilleur espace possible* » et dans un texte postérieur, soulignant l'importance de « (...) *décrire la disposition topologique qui donne à la lecture du texte classique à la fois son tracé et sa liberté.* » [Barthes 84 p.36].

La critique littéraire a peu à peu investi cette notion de topologie. L'une des principales raisons de cet engouement est probablement la pluralité des objets dont une analyse topologique permet de rendre compte. Ainsi quand Tibor Papp<sup>41</sup> évoque la mise en place d'une « *topologie du texte* », celle-ci se construit non plus dans une perspective narratologique mais essentiellement typographique : « (...) *car même en partant des facteurs de la page mallarméenne, pourquoi ne systématiserait-on pas les valeurs graphiques, n'emploierait-on pas un répertoire codé, n'isolerait-on pas des mots (temporairement et spatialement), n'établirait-on pas une topologie du texte ?* »

#### 4.2.2.3. Topologie de l'œuvre.

Emile Noël : « *On pourrait lire la Bible en y cherchant les fronces et les plis ?* »<sup>42</sup>  
René Thom : « *Une sorte de sémiotique topologique de la Bible ?* » [Thom 93 p.57]

Narratologie, typographie, problématique de genres, sémiologie comme dans l'exergue ci-dessus, il n'est aucune des formes de critique et d'exégèse qui ne puisse tirer parti d'une approche topologique. C'est l'œuvre toute entière, l'œuvre comme work in progress qui à son tour est analysée par Shklovsky<sup>43</sup> dans cette optique :

« *Une œuvre littéraire est une forme pure. Ce n'est ni une chose ni une matière, mais une relation entre matières. Et comme toute relation, elle a aussi très peu à voir avec la longueur, la largeur ou toute autre dimension. C'est la signification arithmétique de (...) leurs relations qui est importante.* » Cité par [Hivnor & Porush 95]

Enfin, c'est Steiner qui a, à notre connaissance, bouclé l'écheveau complexe des rapports entre littérature et topologie, posant cette dernière comme le chaînon manquant entre la littérature et l'étude des formes culturelles qu'elle produit ou dans lesquelles elle s'inscrit.

« *On peut qualifier de topologiques ces transformations multiples et ces réagencements de rapports entre un prétexte verbal initial et ses réapparitions successives sous d'autres formes, verbales ou non. (...) La topologie (...) a révélé des traits communs et des groupements sous-jacents parmi une foule de fonctions et d'organisations de l'espace apparemment diverses. Parallèlement, certains invariants et constantes charpentent la multiplicité des formes d'expression de notre culture. C'est grâce à eux qu'il est possible (...) de considérer la trame de la culture comme « topologique ». Quelques unes de ces constantes sont spécifiquement verbales. D'autres thématiques. D'autres encore formelles. (...) L'histoire du topos, de l'archétype, du motif, du genre est un lieu commun de la littérature comparée et de la stylistique contemporaines.* » [Steiner 98 p.571]

<sup>41</sup> cité par [Laufer 89 p.206]

<sup>42</sup> deux des sept catastrophes élémentaires dans la théorie de Thom.

<sup>43</sup> Shklovsky V., **Theory of Prose**, Traduction de Benjamin Sher, Archive Press, 1990 (Première édition russe de 1925).

### 4.3. Topologie et hypertexte.

« *Sur le web tout est sur le même plan. (...) Et cependant tout est différencié.* » [Lévy 00a p.161]

#### 4.3.1. Topologies dynamiques de l'organisation hypertextuelle.

Parmi les caractéristiques « invariantes » de l'organisation hypertextuelle, nous avons insisté tout au long de ce travail, sur la variation et l'entrelacement constants des différents niveaux d'échelle qui s'offrent à la navigation et à l'analyse critique, ainsi que sur la topologie sans cesse en mouvement qui caractérise l'agencement de l'ensemble de ses parties. Soit deux « invariants » dont la nature est d'être « changeante ». Cet oxymore n'est cependant qu'apparent si nous envisageons l'organisation hypertextuelle comme un système autopoïétique, c'est-à-dire :

« (...) *organisé comme un réseau de production de composants qui régénèrent continuellement par leurs transformations et leurs interactions le réseau qui les a produits, et qui constituent le système en tant qu'unité concrète dans l'espace où il existe, en spécifiant le domaine topologique où il se réalise comme réseau. (...) Ainsi, une machine autopoïétique est un système homéostatique (ou, mieux encore, à relations stables) dont l'invariant fondamental est sa propre organisation (le réseau de relation qui la définit).* » [Varela 89 p.45]

Ce sont les principes d'organisation – le fait qu'une organisation existe et soit à différents niveaux plus ou moins perceptible – qui constituent l'invariant de tout système hypertextuel, indépendamment de son échelle, de son degré d'ouverture, du sens qu'il véhicule et de sa place au sein de l'ensemble des unités réticulées qui composent le réseau. Un invariant qui ne peut être posé que d'un point de vue critique, à distance de l'expérience temporelle particulière que fonde chaque session de navigation : pour chacune d'entre elles, la topologie dynamique qui caractérise ce système autopoïétique cède le pas à des topographies statiques. Tout discours s'efforçant de fonder une herméneutique de l'organisation hypertextuelle, adopte de fait, cette perspective topologique. A l'inverse – ou plus exactement de manière complémentaire – chaque nouvelle expérience de navigation considérée pour elle-même, ou la mise bout à bout d'un nombre indéterminé de sessions, relève de la perception du géomètre : il s'agit alors de « prendre la mesure » des liens, du niveau de connectivité des unités qu'ils agrègent et du sens qu'il appartiendra alors à chacun de dégager au travers de leurs propriétés individuelles ou contextuelles. Il ne s'agit plus d'envisager seulement le nombre de nœuds, le fait qu'ils soient ou non liés à d'autres, et la propriété collective d'organisation dont ils sont le reflet.

#### 4.3.2. Topographies statiques de la session hypertextuelle.

Si, en adoptant le point de vue du géomètre, nous entrons maintenant dans le cadre d'analyse offert par la session, plusieurs de ces topographies statiques<sup>44</sup> se mettent simultanément en place selon leurs propres modalités.

<sup>44</sup> statiques à un moment donné, et donc reproductibles et non-linéaires (propriétés temporelles de la session).