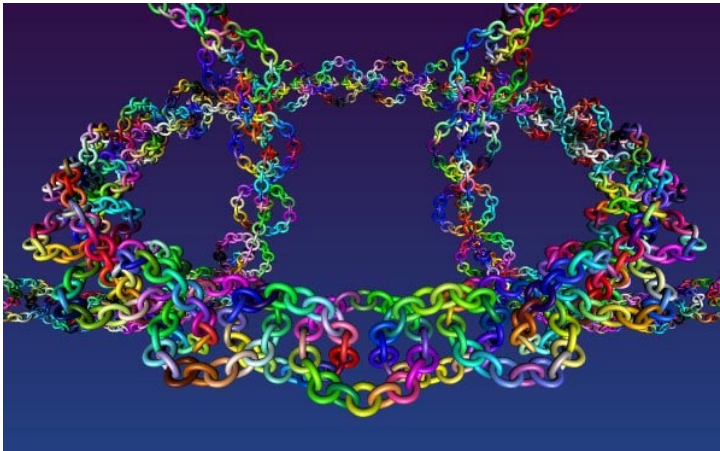


Section C

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]

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 FoRSIC⁵⁸ a occupé une place centrale tout au long de la rédaction de cette thèse et en constitue la partie expérimentale et applicative⁵⁹. 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 connaissance⁶⁰.

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 expérimental⁶¹.

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 »⁶² ;
- la dimension collective : nous verrons de quelle manière s'organisent les agencements collectifs d'énonciation entrant dans le cadre du projet ;

⁵⁸ 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>.

⁵⁹ 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-Carand, 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.

⁶⁰ 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.

⁶¹ 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.

⁶² voir le point 7.6.4. « Invariants procéduraux et déclaratifs. » du chapitre deux.

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 »⁹⁹. Nous préférons le définir comme un espace collaboratif d'apprentissage (CLS)¹⁰⁰. 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.

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

¹⁰⁰ CLS : Collaborative Learning Space.

- 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.

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-agent⁶³, 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 campus⁶⁴ :
 - 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-forme⁶⁵ :
 - 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-K⁶⁶, à déposer des ressources de formation, à décrire des sources de formation, à remplir

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

⁶⁴ 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.

⁶⁵ 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.

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

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 connaissance⁹⁷.

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 exemplaire⁹⁸ : il offre dans une interface unique la possibilité de disposer des

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

⁹⁸ 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 connaissances⁹⁶, 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

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

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 connaissances⁶⁷.

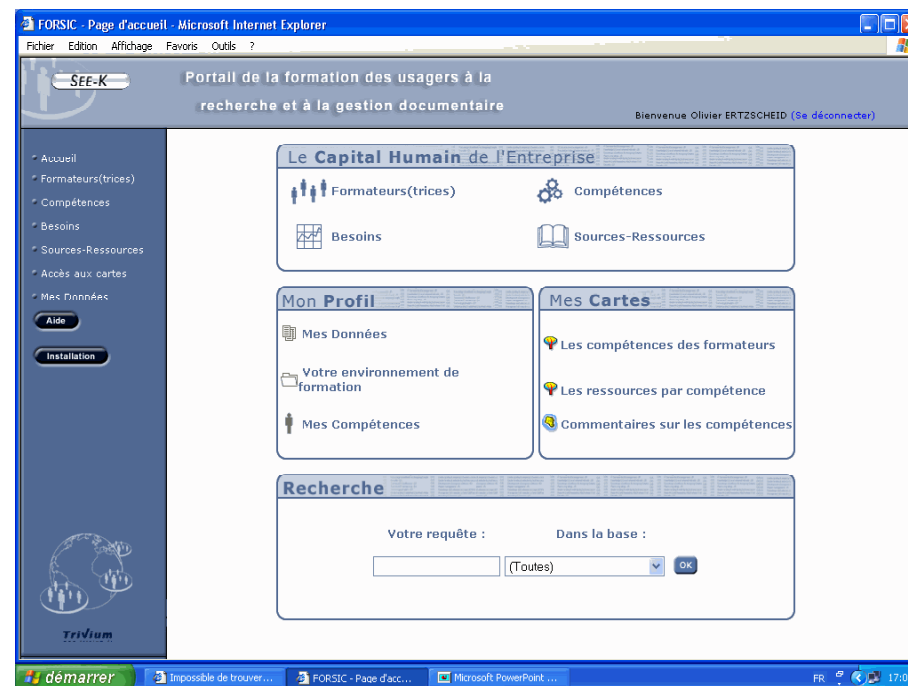


Fig. 16 : Copie d'écran de la plateforme SEE-K.

⁶⁷ voir le point 6.5.3. « Dimension cartographique et arbres de connaissance. »

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 SABRE⁶⁸ 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 document⁶⁹.

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.)

⁶⁸ SABRE : Système Auteur Basé sur la Réutilisation.

⁶⁹ En ce sens, SABRE fonctionne sur les principes du versioning.

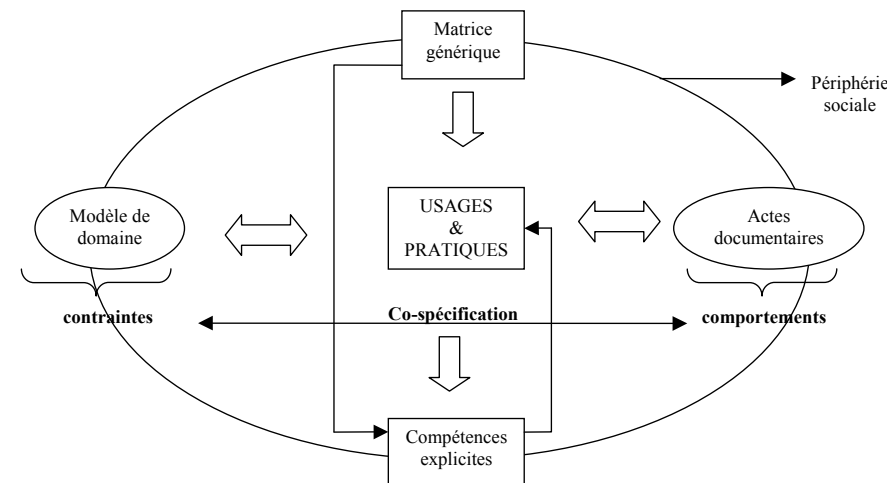


Fig. 23 : FoRSIC, une organisation hypertextuelle.

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 constituante, 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'usage⁹⁴, 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 »⁹⁵ ;
- 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 »).

⁹⁴ celui de l'interface de FoRSIC mais également celui des usages documentaires.

⁹⁵ 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.

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.

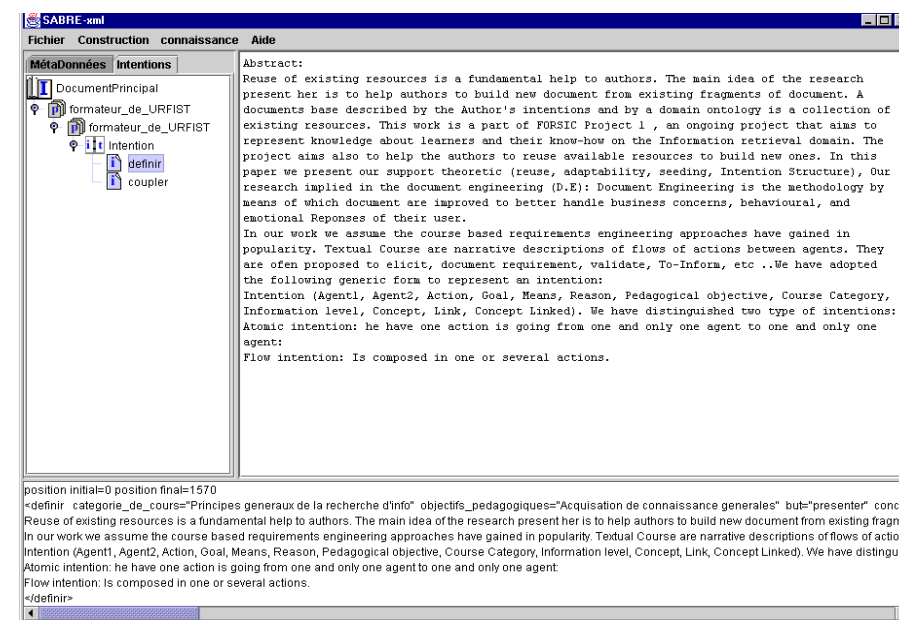


Fig. 17 : Copie d'écran du dispositif SABRE.

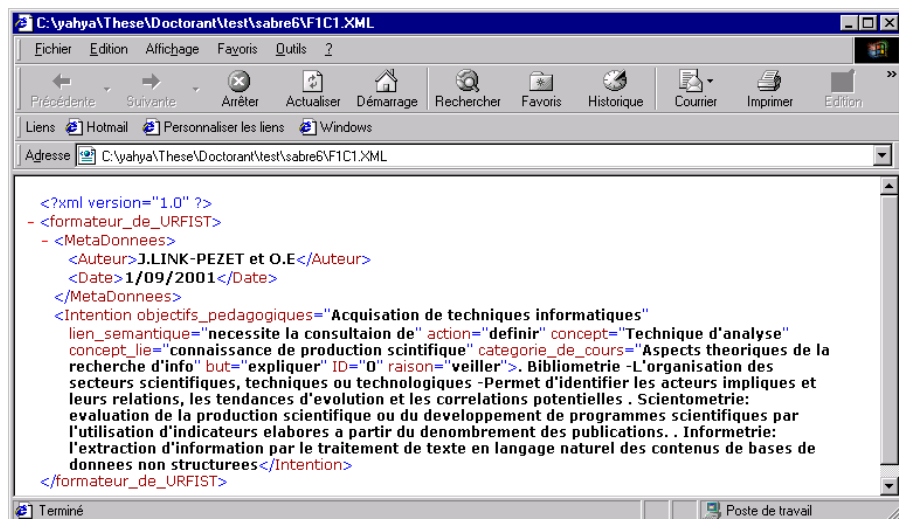


Fig. 18 : SABRE : un exemple de ressource XML produite.

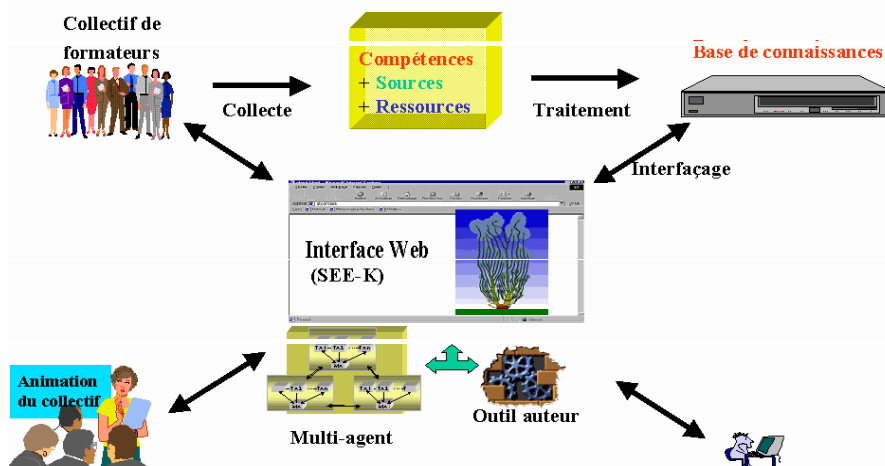


Fig. 19 : Architecture fonctionnelle de FoRSIC.

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 (Délivrable) » 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.* »

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.

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 central⁷⁰. 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,

⁷⁰ 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.

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 retenues⁹³ 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 »).

⁹³ 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.

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 ailleurs⁷¹ 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 ;

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

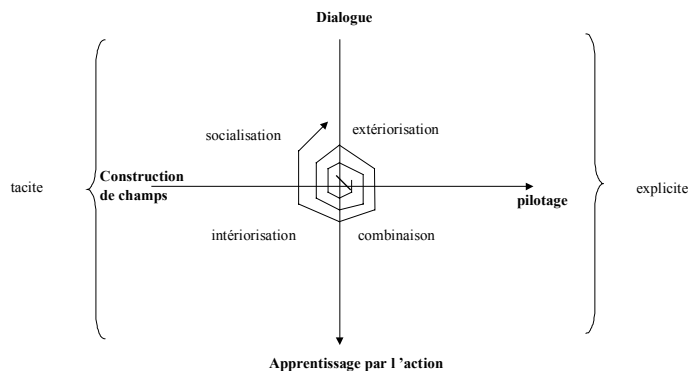


Fig. 20 : La conversion de connaissance d'après [Takeuchi & Nonaka 95].

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

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, notamment⁹² – 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

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

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

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'échange⁷². 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

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

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 maintenus⁹⁰. [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'information⁹¹. 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

⁹⁰ 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.

⁹¹ 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 informatique⁸⁷ ;

- 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 disjointes⁸⁸ ;
- 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 vocabulaire⁸⁹. 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

⁸⁷ [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. » »

⁸⁸ 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.

⁸⁹ 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.

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-coopératives⁷³.

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 éléments⁷⁴, 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]

⁷³ voir le point 7.6.4. « Invariants procéduraux et déclaratifs. » du chapitre deux.

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

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.

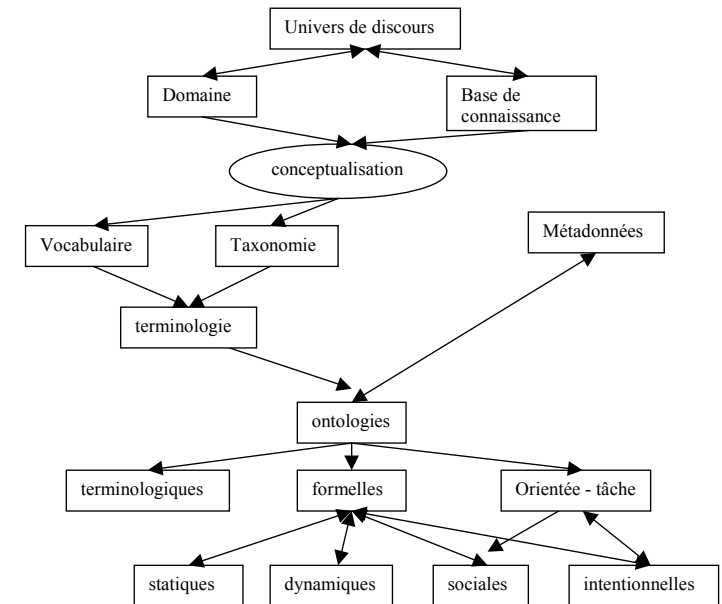


Fig. 22 : Types ontologiques.

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 »⁸⁶. 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.

⁸⁶ [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). »

□ « 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]⁷⁵.

« 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.

⁷⁵ 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.

[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 haut⁷⁶, 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.

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

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 conceptualisation⁸³. 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 connaissance⁸⁴).

« 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 adjectifs génériques permettant de modifier les objets.
 - D'autres concepts spécifiques à la tâche. »

On trouve enfin la distinction opérée par [Jurisica et al. 99]⁸⁵ 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 ;

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

⁸⁴ [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.).

⁸⁵ 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 conceptualisation⁸², 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).

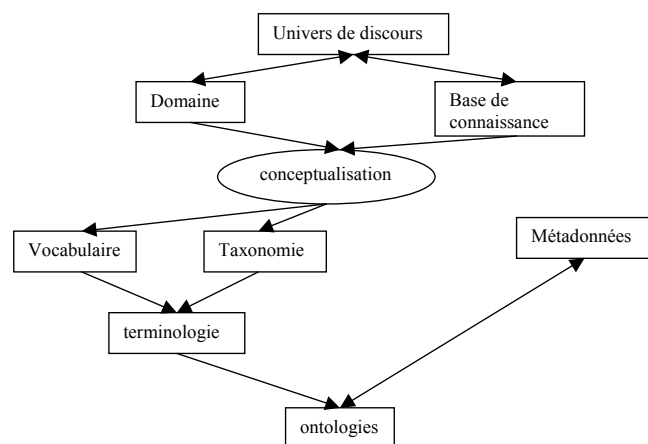


Fig. 21 : Éléments pouvant composer une ontologie.

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.

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

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 UMLS⁷⁷) 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'autres⁷⁸, 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, RDF⁷⁹ ...), de leurs outils de génération (OntoWeb, OntoBroker ...), et de leurs champs d'application⁸⁰.

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

⁷⁷ MESH : Medical Subject Headings, UMLS : Unified Medical Language System.

⁷⁸ 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>.

⁷⁹ 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>).

⁸⁰ 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.

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.

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]⁸¹.

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

⁸¹ 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.