

SemioNet

Spécification, production et implémentation de services
d'information en ligne

*Elisabeth Stockinger de Pablo,
Hammou Fadili, Peter Stockinger*

1. Introduction

Les développements récents des technologies de la communication et de l'information, notamment ceux concernant la mise en réseau des ordinateurs, des logiciels spécialisés et des applications, intéressent plus particulièrement la recherche et l'éducation ainsi que différents métiers parfois très spécialisés dans la gestion, la veille et la diffusion de l'information. Les exemples d'applications les plus connus renvoient aux domaines de l'enseignement à distance et du "campus électronique", de l'édition en ligne, du conseil en ligne ou encore du "travail collectif" (du *groupeware*, en anglais) à distance.

Ces applications se concrétisent sous forme de services et de documents accessibles et utilisables en ligne (sur Internet par exemple). Or, la réalisation de tels services et documents exige une infrastructure physique, technologique et, surtout informatique adéquate mais également des cadres conceptuels et des méthodologies de spécification, de production et de maintenance de ces services et documents .

Le terme "SemioNet" est un néologisme formé à partir des deux termes (anglais) *semiotics* et *network* ou encore à partir de l'expression composée *semiotic networking*. Il renvoie aux problèmes de modélisation, de conception, de spécification, de production et d'implémentation de services et de documents à support électronique (et/ou en ligne), sur leur maintenance ainsi que sur la réutilisation des "scénarios" qui les sous-tendent pour la production d'autres documents et services. Ces problèmes ne sont pas purement d'ordre technologique mais d'ordre sémiotique (ou conceptuel) ; leur résolution mobilise une grande variété de réponses mais ne concerne habituellement que telle ou telle facette de la problématique

envisagée dans sa généralité. Différemment, notre point de vue propose une “vision” ou une “approche” sémiotique [Greimas, Courtés, 1979 ; Eco, 1972] du document qui pourrait constituer un cadre d’intégration à la fois cohérent et souple des différentes méthodologies locales existantes et ainsi former la base d’une méthodologie systématique et générale pouvant guider les activités de la spécification, production, maintenance et réutilisation de documents et services à support électronique. Cependant, le propos de cet article n’est pas de problématiser le cadre ou encore la “vision sémiotique” du document (*i. e.* du *texte* au sens le plus large du terme). Il existe en ce domaine des ouvrages d’introduction qui ont fait leurs preuves (cf., par exemple [Eco, 1972]). Notre ambition est d’introduire, à partir d’un exemple concret de réalisation d’un site *Web*, les différents aspects que devrait traiter une méthodologie systématique d’inspiration sémiotique et d’en proposer une première esquisse.

2. La problématique générale

2. 1. La bibliothèque numérique

Le développement de produits (services, documents, applications multimédias) pour la valorisation et la communication d’un patrimoine culturel ou scientifique fait partie d’une problématique plus générale qui est celle de la mise en place de “bibliothèques numériques” (*digital libraries*, en anglais) [Adam, Bhargava, Yesha, eds., 1995].

Une “bibliothèque numérique”, entendue d’une manière simple et très générale, est constituée par une quantité importante de données — de documents — multimédias (*i. e.* de textes, d’images statiques, de vidéos, etc.). Ces données en constituent le fonds et sont stockées dans plusieurs bases de données qui peuvent se trouver à des endroits physiques éloignés.

Autour de ce fonds documentaire se met en place un ensemble de services particuliers qui doivent assurer la bonne gestion du fonds ainsi que son exploitation optimale par les utilisateurs de la bibliothèque. Parmi ces services, on compte les services standards d’une bibliothèque ou d’un centre de documentation et d’information “traditionnel” tels que l’aide à la recherche d’un document pertinent, le prêt d’un document, la constitution de dossiers à thème, la mise à jour et la maintenance du fonds documentaire qui garantissent une bonne qualité des documents disponibles, et ainsi de suite (cf. par exemple [Taesch-Wahlen, 1997]). Mais, à côté de ces services émergent, dans le contexte des bibliothèques électroniques, des services nouveaux qui exigent des compétences particulières constituant souvent des métiers spécialisés.

Ces services concernent, à titre d'exemple, la veille d'information sur réseau, l'évaluation des documents et des informations disponibles sur réseau, la production de documents à partir de documents constituant le fonds d'une bibliothèque numérique, la rédaction à distance de documents, la traduction et l'adaptation à distance de documents, la mise en place d'environnements de coopération (scientifique ou autre) visant la réalisation collective d'un produit ou encore la formation à distance.

Une bibliothèque numérique ne doit pas être vue exclusivement sous l'aspect restreint de "bibliothèque" ou de "centre de services d'information professionnelle". Elle prend place dans des structures intégrées de services d'information, de documentation, d'édition, de conseil et de formation — structures qui peuvent être réalisées aussi bien sous forme de sociétés spécialisées que sous forme d'environnements technologiques (*i. e.* de logiciels) appropriés à ces tâches à la fois diversifiées et complexes faisant partie de l' "industrie de l'information".

La constitution, la gestion et le développement de bibliothèques numériques, et des services qui gravitent autour d'une telle bibliothèque, peuvent être répartis en un ensemble d'activités typiques. Ces dernières constituent, en quelque sorte, le noyau conceptuel de tout projet de gestion électronique d'information et de documentation (cf. également à ce propos [Dupoirier, 1995] ainsi les contributions de Chaumier, Sutter et Turner dans le [D. E. I. D., 1997]). Nous les classons comme suit :

- Recherche de l'information pertinente et collecte de documents contenant l'information recherchée.
- Acquisition des informations et documents collectés.
- Indexation de documents saisis et numérisés.
- Stockage des documents indexés.
- Rédaction, traitement et production de nouveaux documents.
- Edition (au sens de "fabrication") de documents.
- Consultation d'un fonds documentaire.
- Communication et diffusion d'un fonds documentaire.
- Archivage de documents possédant une "valeur historique".

Chacune de ces activités se compose d'un ensemble de tâches plus spécifiques dont la description est particulièrement importante aussi bien pour la bonne gestion du "flux du travail" et de la coordination des différentes actions, que pour la mise en place de standards, c'est-à-dire de normes techniques qui constituent la "colonne vertébrale" des logiciels et des autres outils spécialisés de travail.

Faute de place, nous ne discuterons pas ici ces activités d'une façon détaillée. Nous nous contenterons d'une brève présentation de certaines d'entre elles¹.

¹La description et la problématisation détaillées des différentes activités citées se trouvent dans la version électronique de Semionet sur le site de l'Equipe Sémiotique Cognitive et Nouveaux Médias — ESCoM (<http://vhf.msh-paris.fr/Escom>).

2. 2. Recherche et collecte d'informations

L'activité concernant la recherche et la collecte de documents doit répondre à un objectif à satisfaire. Un objectif d'information peut être de nature simple, claire et ponctuelle comme par exemple celui de la requête d'un ouvrage précis ou encore d'un ensemble de documents qui proposera des informations sur un thème choisi. Mais il ne s'agit ici que d'un cas limite. "Objectif d'information" peut signifier constitution de dossiers ou de documents complexes, réalisation de publications/périodiques, mises à jour d'un fonds documentaire par acquisition de nouveaux documents, et ainsi de suite.

Aussi bien dans sa forme traditionnelle, que comme activité s'exerçant sur Internet (ou sur d'autres réseaux), la recherche et la collecte d'informations consistent en l'identification de "sources d'information" parmi lesquelles on peut espérer obtenir des informations non seulement pertinentes par rapport à un objectif donné, mais également de "bonne qualité", des "informations à forte valeur ajoutée", etc (cf. notamment [Pomart, 1997, p. 591-593]).

Sous le terme de "source d'information" nous réunissons des documents de toutes sortes mais également les établissements tels que les centres de documentation et/ou de recherche, les bibliothèques, ou encore certaines personnes privilégiées telles que les experts, les spécialistes, etc.

En tant qu'activité en ligne (se réalisant, par exemple, sur Internet), elle s'appuie sur des interfaces de consultation standard, comme le Navigator de Netscape ou l'Explorer de Microsoft, ainsi que sur des moteurs de recherche (fonctionnant par mots clés, domaines, zones géographiques, ...) tels que Yahoo, Excite ou Alta Vista (en ce qui concerne le monde anglophone) Locace ou Francite (en ce qui concerne le monde francophone), Yellowweb Europe (en ce qui concerne plus particulièrement les sites de la CE), ALIWEB (en ce qui concerne les sites), ARCHIE (en ce qui concerne le monde des logiciels), etc. Il existe, en effet, une grande variété de moteurs de recherche spécialisés pour tel ou tel type de documents, tel ou tel domaine d'expertise, telle ou telle langue, etc.

Vu la très grande quantité d'information disponible sur un réseau, la recherche, l'identification et la collecte d'informations pertinentes et à forte valeur ajoutée constituent progressivement un véritable métier de "veilleur d'information sur réseau", d' "évaluateur" ou encore d' "expert en évaluation". Il va de soi que l'expertise sémiotique joue un rôle important, voire central à condition que l'on maîtrise à la fois le domaine d'expertise (le *Sachgebiet* en allemand) et l'environnement technologique.

2. 3. Saisie des documents

Dans le cas de la gestion électronique de documents, l'activité concernant l'acquisition des informations et des documents collectés est

définie à partir de :

- la saisie “manuelle” des informations à l'aide de logiciels de traitement de texte,
- la numérisation des données collectées (à l'aide d'outils d'OCR (*i. e. Optical Character Recognition*), de scanners, d'outils de numérisation photo, vidéo, son, ...),
- le transfert de fichiers informatiques par réseau et la conversion des fichiers externes (conversion de formats de fichiers, conversion de la mise en page à l'aide d'éditeurs SGML, HTML, ... spécialisés).

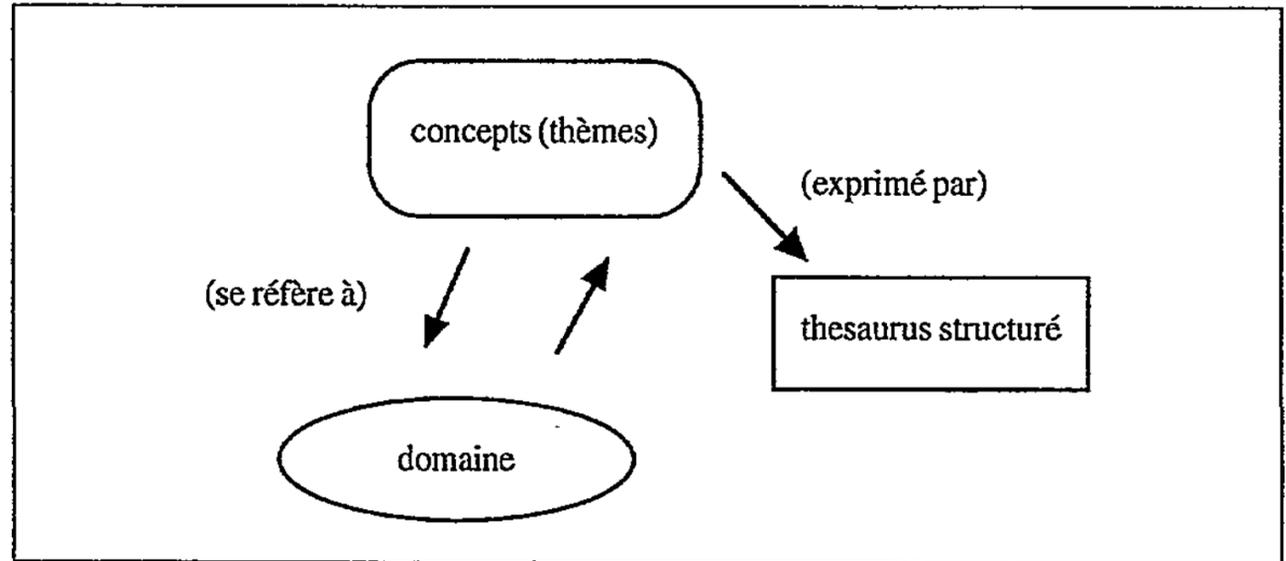
L'acquisition des documents est réglée par un ensemble de standards technologiques et de logiciels spécialisés qui facilitent et optimisent ce processus et qui le rendent accessible aux non-spécialistes.

2. 4. Construction de thesauri et d'index

L'activité concernant l'indexation de documents saisis repose sur la création d'un index et/ou sur l'utilisation (adaptée) d'un index déjà existant. Les index utilisés peuvent être des index standards ou quasi standards justifiés par des descripteurs organisés dans des thesauri plus ou moins généraux ou, au contraire, spécialisés selon des domaines d'expertises particuliers. Les standards ISO 2788 et ISO 5964 proposent des directives pour l'élaboration et la maintenance des thesauri (thématiques, monolingues, multilingues, ...) et pour une introduction excellente dans la problématique de la construction et maintenance d'un thesaurus, on peut se reporter à [Aitchison, Gilchrist, 1992]. Bien entendu, le recours à des standards n'est pas obligatoire — il existe toute sorte d'index *in-house* (*i. e.* utilisé par un seul organisme) sans parler des index personnels dont le créateur est aussi l'auteur d'une application.

Cela étant, l'existence de ces différents types d'index rend difficile, sinon pratiquement impossible, la recherche et l'exploration d'informations qui se trouvent stockées dans différentes bases de données répondant chacune à son propre index. D'où les tentatives de création (semi-)automatique d'index sous forme de bases terminologiques. Ces tentatives se basent sur les descriptions linguistiques (morpho-syntaxiques, en particulier) des “termes”, mais il faut bien reconnaître que le réel travail d'indexeur (*i. e.* la détermination d'une liste de mots-clés autorisés, la constitution d'un thesaurus à prétention thématique) ne peut être pris en considération par une telle approche dépourvue d'une sémantique lexicale et notamment de connaissances expertes.

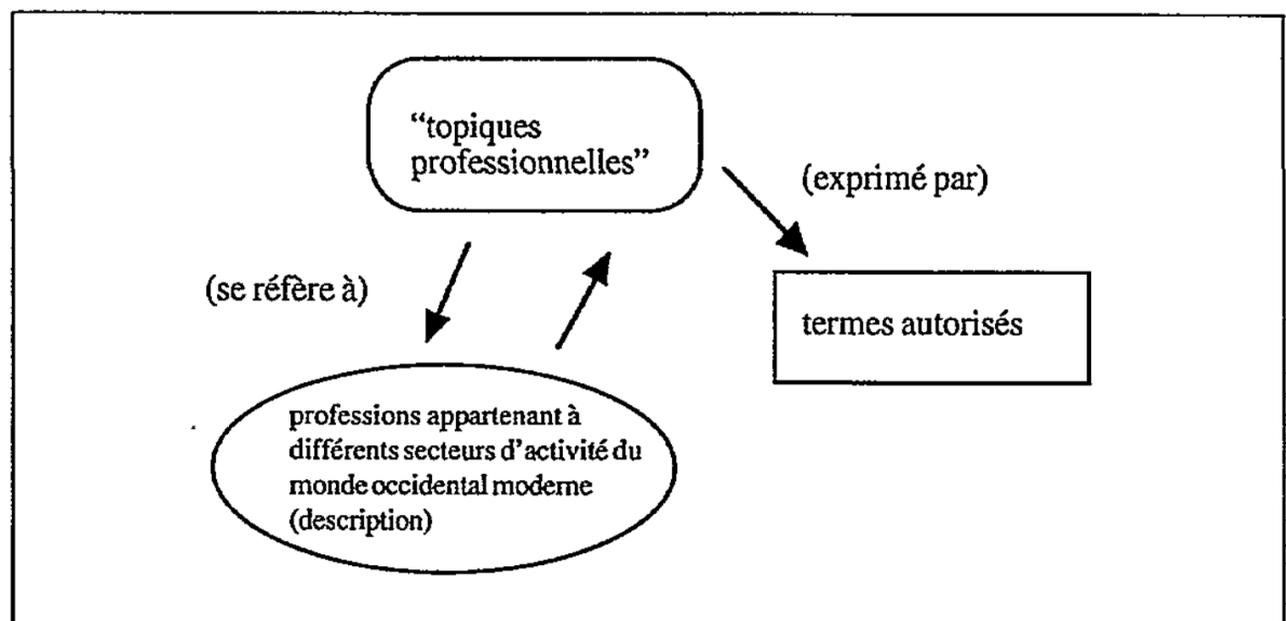
Le graphique ci-après montre schématiquement la problématique générale de la construction d'un thesaurus — problématique qui consiste à mettre en place un vocabulaire contrôlé (des descripteurs) exprimant une théorie, une “vision” d'un domaine de référence.



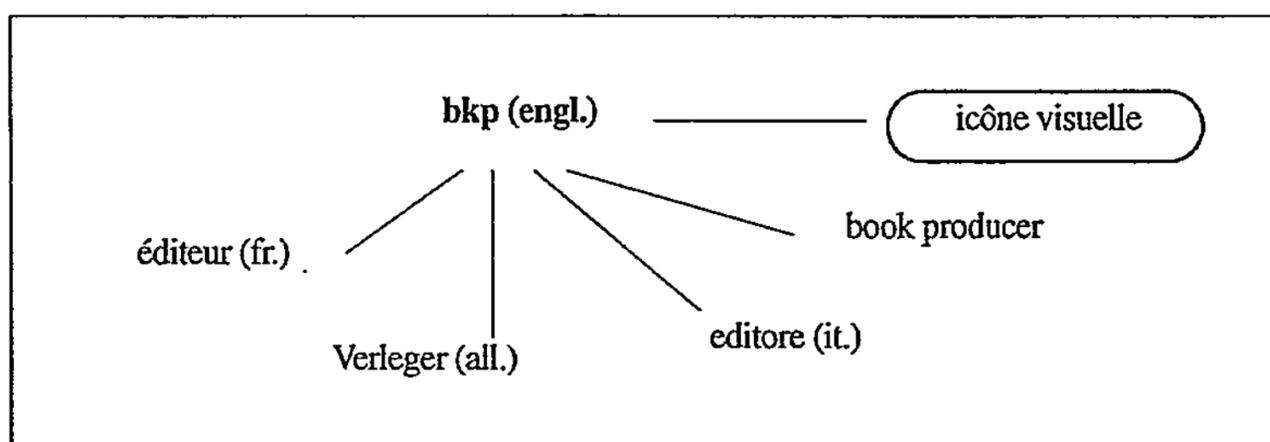
Prenons un exemple. La *Library of Congress* utilise un thesaurus pour l'indexation et la recherche de documents proposant des informations relatives à des activités professionnelles. Ce thesaurus, dont on trouvera ci-dessous un extrait sous forme d'une simple liste de descripteurs, est utilisé par un très grand nombre de bibliothèques aux USA et ailleurs, ce qui facilite la recherche dans les bases de données textuelles distribuées sur le continent nord-américain.

act	Actor
arc	Architect
aus	Author of screenplay
aut	Author
bkp	Book producer
chr	Choreographer
cil	Calligrapher
cmp	Composer
col	Collector
cph	Copyright holder
cst	Costume designer
dnc	Dancer
edt	Editor....

Le schéma ci-dessus, restreint au domaine "activités professionnelles", se présente ainsi :



Dans la liste présentée ci-dessus, on trouve la présence d'expressions notationnelles ou notations (arc, aus, aut...) et d'expressions nominales en langue anglaise (Architect, Author of screenplay, Author...), or la distinction entre le système de notation (indépendant de la langue naturelle) et le code d'expression ouvre la voie vers la construction de thesauri multilingues mais aussi "multilangages", c'est-à-dire des thesauri comportant aussi bien des expressions "linguistiques" que visuelles (par exemple des icônes) ou encore des thesauri comportant différents descripteurs qui peuvent être "réunis" sous un symbole notationnel (parce qu'on considère qu'ils sont, sémantiquement, plus ou moins équivalents). Cf. à ce propos l'exemple (élémentaire) suivant :



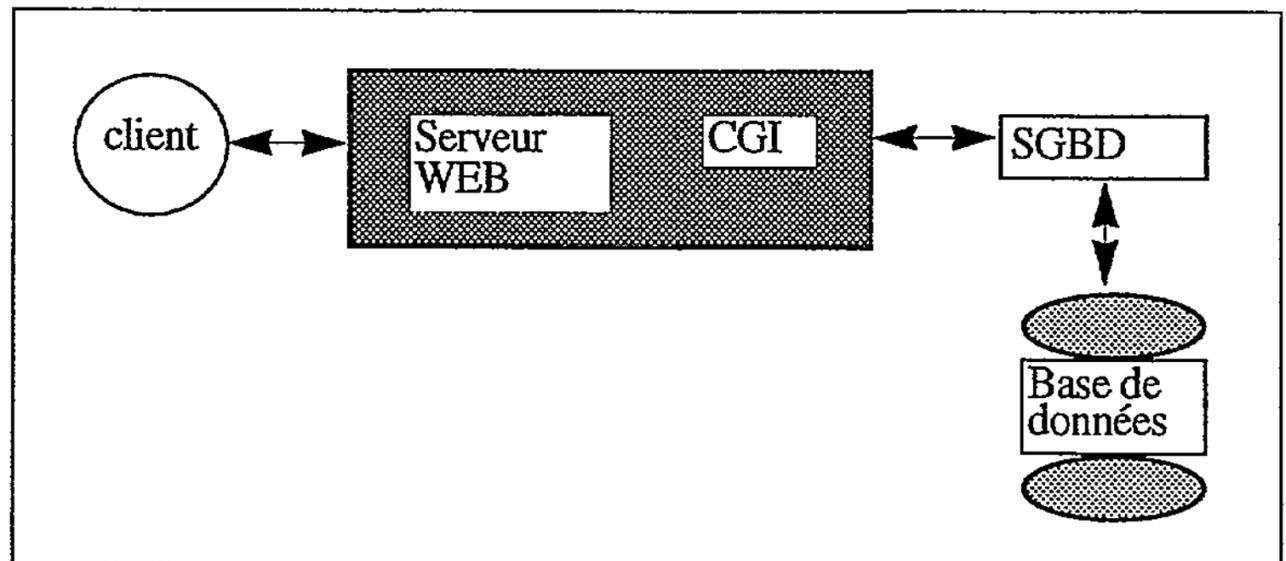
Le système notationnel, bien sûr, constitue l'expression (l' "alphabet") de la théorie descriptive du domaine, c'est-à-dire de la topique ou encore de la configuration thématique à l'aide de laquelle est "vu" et "interprété" un domaine comme celui des activités professionnelles.

On voit donc que la problématique de la constitution de thesauri rejoint des questions relatives à ce que l'on a appelé la modélisation et la représentation des connaissances [Sowa, 1984 ; Pfeiffer, Nagle, 1993] qui ne peuvent être prises en charge que très partiellement par une approche linguistique ou lexicale.

2. 5. Stockage des documents

L'activité concernant le stockage des documents indexés renvoie au choix entre différents modes de stockage (sous forme de fichiers organisés dans des répertoires ou sous forme de données proprement dites dans des bases de données) ainsi qu'à la sélection du matériel de stockage approprié. D'autres tâches plus spécifiques concernent l'utilisation des techniques de compression/décompression des fichiers volumineux (ceux contenant notamment des documents vidéos) ainsi que la mise en place des procédures garantissant la sécurité des fichiers et des données (sécurité s'entend aussi bien au sens de l'intégrité physique des fichiers et des données qu'au sens — juridique — des droits d'utilisation, d'exploitation ou encore de modification des documents stockés).

Le stockage des données présuppose, d'une part la création de schémas conceptuels (de "schémas directeurs" ou encore de "méta-données") suivant lesquels sont classées et stockées les différentes données documentaires. La mise en place de tels schémas renvoie aux procédures visant la définition de topiques ou configurations thématiques à l'aide desquels un domaine particulier est appréhendé et servent à définir les descripteurs dans un thesaurus. Ces schémas conceptuels sont également utiles dans la conception des Bases de Données Objets ou Relationnels qui s'imposent dans le monde des technologies et représentent des solutions de stockage d'informations. Ce type de structure permet d'accéder, avec une performance considérable, aux données et détient une modularité évolutive pour les environnements distribués utilisant des couplages entre nouvelles technologies comme CORBA, DCOM/OLE... etc. et des SGBDOR comme ORACLE, SQL server, Informix, etc. où les bases fournissent les données aux programmes qui génèrent dynamiquement des documents en ligne suivant le choix des utilisateurs. Ces programmes s'appellent *scripts* ou CGI (*Common Gateway Interface*) et sont écrits dans des langages spécialement conçus pour la programmation Internet comme Java, Perl, JavaScript, ActiveX...



2. 6. Rédaction, production et édition de documents

L'activité concernant la rédaction, le traitement et la production de documents constitue certainement le domaine le plus intéressant pour une sémiotique documentaire. Cette activité peut et doit être décomposée en plusieurs tâches spécifiques telles que :

- la spécification conceptuelle et technique d'un scénario définissant la composition structurale d'un document (d'une application) à produire,
- la sélection des documents stockés (sous forme de fichiers ou de données),

- l'assemblage progressif des documents suivant le scénario et à l'aide de logiciels spécialisés (ex. : Director, Authorware),
- le traitement local de chaque fichier ou donnée sélectionnée (à l'aide de logiciels spécialisés tels que Photoshop, Premiere, Illustrator, SoundEdit, etc.),
- la réalisation de fonctions plus évoluées à l'aide d'une programmation (ex. : mise en place de *scripts* à l'aide du langage Lingo pour Director ; réalisation de *scripts* CGI pour des applications en ligne),
- la réalisation de tests, d'évaluations et de corrections du document réalisé.

Nous discuterons plus en détail la notion de scénario sémiotique dans le paragraphe 5. de cet article. Sans vouloir problématiser davantage cette activité, ajoutons seulement que la maîtrise de celle-ci présuppose tout de même des connaissances techniques (technologiques) dans la mesure où elle ne peut être réalisée que grâce au concours d'un environnement technologique assez diversifié. Cet environnement est constitué, entre autres, par des logiciels et autres outils spécialisés tels que :

- 1) des générateurs d'applications multimédias — de documents ou services — en ligne ou hors-ligne (exemples logiciels : Director, Authorware),
- 2) des gestionnaires de bases de données documentaires (Infomix, Oracle, SQL, ...),
- 3) des outils pour le traitement des composants de présentation d'un document (exemple logiciel : Illustrator, X Res),
- 4) des outils pour le traitement de données typographiques et textuelles (exemples logiciels : XPress, PageMaker ...),
- 5) des outils pour le traitement de composants graphiques et photographiques dans une application (Exemples logiciels : Photoshop, XRes, ...),
- 6) des outils pour le traitement des composants vidéo dans une application (Exemple logiciel : Premiere, Avid),
- 7) des outils pour le traitement des composants son (Exemple logiciel : SoundEdit),
- 8) des outils pour le traitement des objets en 3D dans une application (Exemples logiciels : StrataStudio, Extreme 3D, Dimension...),
- 9) des éditeurs *Web* pour des applications en ligne (exemples logiciels : FrontPage et InternetStudio, PageMill, SiteMill...),

10) des éditeurs structurés pour la réutilisation de différents composants d'une application ou le portage d'un composant particulier (Exemples logiciels : famille de logiciels DynaTools).

Habituellement, on distingue les tâches qui concernent la production (intellectuelle, infographique, informatique, ...) d'une application multimédia des tâches qui concernent l'activité de l'édition *stricto sensu*, c'est-à-dire de sa fabrication consistant dans la réalisation d'une maquette CD-ROM, d'un document papier ou encore d'un document vidéo.

D'une manière plus générale, la notion d' "édition électronique" recouvre l'ensemble des activités particulières citées ci-dessus à savoir : l'acquisition de documents, leur indexation et stockage, le traitement de documents existants, la production d'un nouveau document (d'une nouvelle application) ainsi que la fabrication, c'est-à-dire la production d'un master, le pressage d'un CD-ROM ou encore l'impression d'un document à support papier.

Il faut bien être conscient que l'activité de l'édition électronique, entendue au sens large du terme, possède ses propres particularités difficilement intégrables dans les chaînes de production des documents traditionnels. Ces particularités renvoient à :

- la production, la réalisation (ainsi que le stockage et la diffusion) de versions personnalisées, de versions adaptées à un besoin spécifique, de versions adaptées au niveau de connaissances d'un client, etc. ;
- la production d'une version particulière à partir d'un ensemble de versions de documents déjà existants, de même support ou de supports différents et stockés physiquement dans un même endroit ou dans plusieurs endroits différents ;
- l'intégration d'une dimension interactive dans le processus de production même (*i. e.* la possibilité pour un client, de pouvoir modifier certaines décisions structurant le processus de la production-réalisation) ;
- la réalisation d'une version concrète sur différents supports (par exemple : sur support papier et sur support CD-ROM) ;
- la spécification, la production et la réalisation d'une version concrète/de plusieurs versions concrètes par plusieurs auteurs travaillant à distance ;
- la modification ou encore la mise à jour "ciblée", "locale" d'une version concrète, c'est-à-dire la production de nouvelles versions légèrement différentes de celle d'origine ;
- l'archivage, le stockage et la diffusion de versions concrètes tenant compte des conditions d'accès, des conditions d'utilisation, des conditions de modification d'une version concrète ;

— et plus particulièrement pour édition électronique en ligne : la prise en compte et l'intégration des aspects de sécurité, de confidentialité ou encore de paiement.

2. 7. Consultation d'un fonds de documents

L'activité concernant la consultation du fonds d'une bibliothèque numérique ressemble à différents égards à l'activité de la recherche et collecte de documents dans la mesure où celle-ci repose sur l'utilisation de butineurs (*browsers*, en anglais) ou de moteurs de recherche à mots-clés. Néanmoins, ce qui est propre à cette activité, c'est, avant tout, le fait que les interfaces et programmes de recherche, d'accès et de navigation d'information aient été conçus et réalisés afin de satisfaire des objectifs d'information préalablement définis (il est à noter que la définition de ces objectifs peut couvrir une grande variété d'objectifs spécifiques ...).

L'accès et la navigation dans un fonds documentaire d'une application en ligne sont spécifiés sous forme de *scripts* (*i. e.* de mini-scénarios typiques). Ces *scripts* sont comparables aux "guides de lecture" tels que les sommaires, les parcours de lecture, les renvois, etc. L'élaboration des *scripts* d'accès et de navigation présuppose un travail important de description et de modélisation thématiques d'un corpus (d'une partie représentative du fonds) documentaire devant déboucher sur une systématique des configurations qui représentent, sur un certain niveau de granularité, le contenu du fonds en question.

Bien entendu, l'accès et la navigation simple, sans conditions, ne constitue qu'un cas limite de l'exploration d'un fonds documentaire. Des formes plus sophistiquées concernent, par exemple, l'exploration personnalisée (adaptée au "profil" d'un utilisateur — à ses connaissances, ses habitudes, à son statut et sa position) ou l'exploration ludique et/ou éducative (correspondant à un parcours didactique, à l'établissement d'un "curriculum" ou à un jeu à caractère éducatif).

3. Le Serveur expérimental *Orient*

3. 1. Remarques introductives

Comme nous l'avons dit précédemment, nous souhaitons introduire progressivement et discuter le cadre méthodologique devant guider la spécification et la réalisation de services et documents à support électronique appelé SemioNet et ce, à l'aide d'un exemple concret.

Notre exemple porte sur la spécification, l'implémentation et la validation d'un ensemble de services d'information pour la valorisation d'un patrimoine culturel et/ou scientifique. Ces services d'information

s'organisent et évoluent autour d'un serveur expérimental central du nom d'*Orient* (URL : <http://vhf.msh-paris.fr>). Les différents types de services concernés se répartissent comme suit :

- services de consultation d'informations générales ou à thèmes,
- services de publication et d'édition électronique en ligne,
- services de travail et de coopération à distance,
- services d'enseignement et de formation à distance,
- services de commerce électronique.

Le serveur en question doit gérer un réseau de sites spécialisés qui se différencient en plusieurs types principaux :

- des sites pour équipes et centres de recherche en sciences humaines ;
- des sites "à thème" dont le propos est de réunir, autour d'une thématique donnée, des documents, des applications et des sites *Web* pertinents, d'organiser un espace de travail et d'échange pour une communauté "virtuelle" composée de spécialistes mais aussi d'utilisateurs intéressés par la thématique proposée ;
- des sites pour associations travaillant dans la valorisation et la diffusion d'un patrimoine scientifique et culturel ;
- des sites pour éditeurs académiques et/ou spécialisés .

La conception et le développement de ces services font partie du projet Européen *An Environment for a Virtual Hypermedia Factory* (VHF) du programme ESPRIT IV (N° 22251). Le projet VHF porte, plus généralement, sur la conception et le développement d'outils (logiciels), de services d'information et d'édition électronique multimédia en ligne dans le domaine de la gestion du patrimoine culturel (au sens large) d'une institution (académique ou non, privée ou publique).

L'ensemble du projet est coordonné par Olivetti Ricerca en Italie. Le consortium se compose de différents partenaires des pays suivants : Italie, Autriche, Grèce, France qui appartiennent tant au monde académique qu'au monde de l'industrie et de l'édition. La partie française est assurée par la Maison des Sciences de l'Homme (MSH) qui a comme partenaire associé l'Institut National des Langues et Civilisations Orientales. Pour plus d'informations concernant le projet VHF, on peut consulter l'URL (URL : <http://vhf.msh-paris.fr>).

3. 2. Les services et sites en cours de développement

La place nous manque pour présenter le déroulement et le résultat de l'enquête qui a précédé l'identification des différents types de services

pouvant satisfaire des besoins particuliers d'information. Notons seulement que l'enquête a été menée auprès d'un panel représentatif d'utilisateurs institutionnels appartenant aux secteurs suivants : bibliothèques et centres d'information et de documentation, maisons d'édition et librairies, médias (presse audiovisuelle), équipes et centres de recherche, associations, ministères, opérateurs touristiques.

L'objectif de l'enquête a été de préciser d'une manière aussi explicite que possible :

1) les types de services et produits d'information à mettre en place afin de pouvoir satisfaire un besoin d'information émanant d'un utilisateur institutionnel (en tenant compte, plus particulièrement des types d'informations, des types de médias et des genres documentaires les plus demandés) ;

2) les types de sites spécialisés devant être hébergés par le serveur (en tenant compte plus particulièrement des "composants" principaux et de l'organisation interne de ces sites ainsi que de leurs fonctions dans la "vie" ou le "travail" d'un utilisateur institutionnel).

Dans l'enquête ont été distingué deux cibles : une première visant à déterminer l'existant (*i. e.* les services d'information et de communication existants), la seconde visant à déterminer les besoins ou désirs devant être pris en considération lors de l'identification et de la spécification des services et produits à développer dans le cadre du projet Européen cité.

L'enquête elle-même a été réalisée sous forme de réunions de groupe et d'interviews extensifs organisés d'une manière directive, c'est-à-dire suivant un questionnaire préétabli (une version simplifiée est consultable sur le serveur *Orient*). A partir des résultats obtenus, cinq types principaux de services d'information — dont l'objectif principal est, répétons-le, la valorisation d'un patrimoine culturel et/ou scientifique — ont été identifiés, sachant qu'un type principal de service d'information peut recouvrir plusieurs services d'information plus spécialisés.

D'une manière analogique à la bibliothèque traditionnelle, un service d'information *stricto sensu* consiste en un fonds documentaire (textes, images, vidéos, cartes, dessins, ...) classé et stocké soit dans un serveur expérimental et exploitable à l'aide d'un butineur (*browser*) thématique ou d'un moteur de recherche par mots-clés. Cela étant, un service d'information *latu sensu* inclut, outre le composant bibliothécaire, divers autres composants tels que la production et l'édition électroniques, la coopération à distance, la messagerie électronique, etc.

Les différents services constituent un ensemble de "briques" à l'aide desquelles est bâti le serveur. L'architecture de ces services reste ouverte dans la mesure où de nouveaux services peuvent être spécifiés et introduits dans l'architecture existante.

Voici la liste des principaux types de services actuellement en cours de réalisation :

— Services d'informations à visée encyclopédique et à thème

- INFORMATIONS ENCYCLOPÉDIQUES
- SITES À THEMES / DOSSIER THÉMATIQUE

— Services «Travail & Conseil»

- RECHERCHES ET OFFRES D'EMPLOIS

— Service «Voyage & Tourisme»

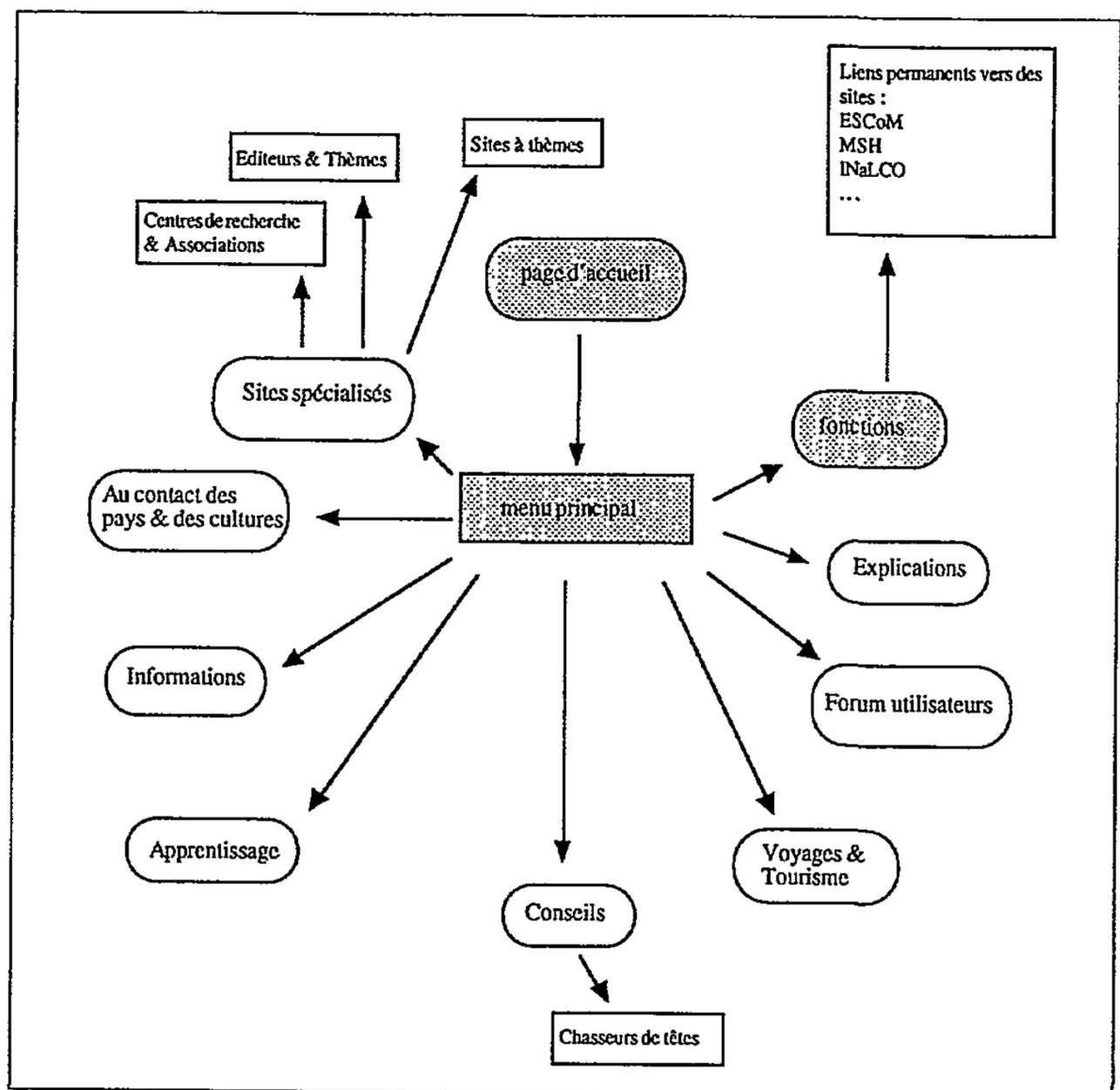
- VALISE POUR TOURISTES

— Services «Au Contact des Pays et Cultures»

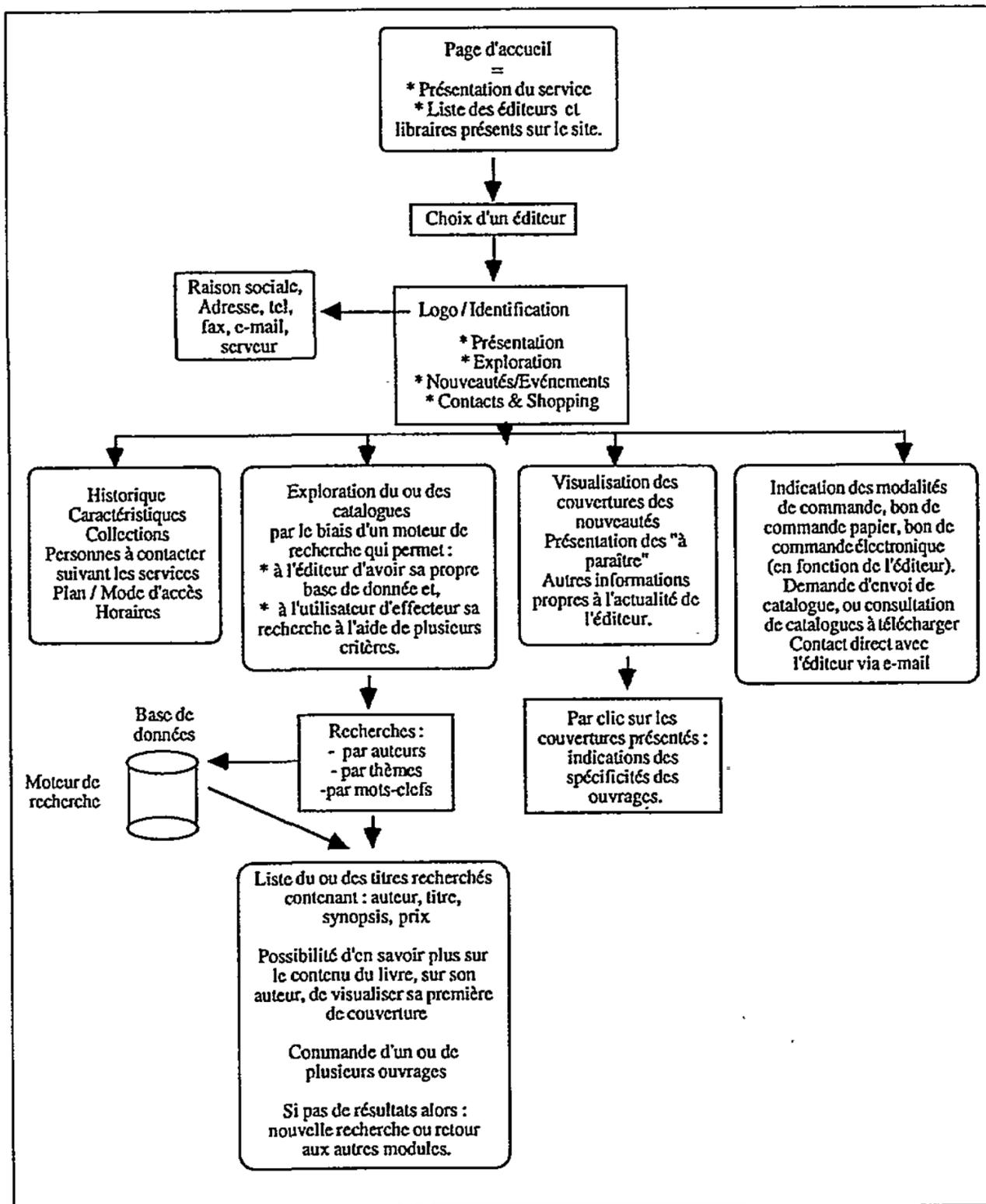
- ASIE À PARIS
- AFRIQUE À PARIS

— Service «Professions et Métiers»

- Les ÉDITEURS CULTURELS & ACADÉMIQUES EN LIGNE
- ORGANISMES D'ÉDUCATION ET DE RECHERCHE
- ASSOCIATIONS / SITES À THEMES



Architecture actuelle du serveur culturel *Orient*



Organigramme du site spécialisé *Les Éditeurs en ligne*

La plupart des services cités ci-dessus peuvent être très facilement redéfinis au sens de types spécialisés de sites *Web*, car ils sont parfaitement indépendants les uns des autres. Par exemple, le service des *Éditeurs en ligne*, qui est un composant du serveur culturel, peut être éclaté en une série de sites indépendants dont chacun correspond à un éditeur particulier (v. le schéma ci-dessus). Il aura ainsi un site propre aux Éditions MSH, aux éditions You-Feng, à CID-Diffusion, etc. De même, le service *Sites à Thèmes / Dossiers Thématiques* peut être transformé et éclaté en sites spécialisés dont l'une des tâches est la réunion et la diffusion d'informations concernant un thème spécifique (une problématique de recherche).

En ce sens, le serveur culturel appelé *Orient* (v. le schéma ci-contre) tend à devenir une sorte de *centre électronique* hébergeant et gérant des

sites spécialisés, des services et ressources d'information pour la valorisation et la diffusion d'un patrimoine culturel ou scientifique — centre électronique organisé comme un réseau (à la fois physique, informatique et conceptuel ou thématique), dont l'exploitation optimale dépend d' "agents" spécialisés, d'applications spécialisées qui, d'une manière analogue à des opérateurs humains, accomplissent des tâches faisant partie des activités "standards" d'un tel centre (cf. à ce propos le paragraphe 4).

3. 3. Vers un système d'information plus "intelligent"

Il faut noter que les différents services que nous venons de citer peuvent être considérés comme des applications locales d'un système *desktop* de recherche et de diffusion d'informations pour la valorisation d'un patrimoine culturel et/ou scientifique — système qui comporte les caractéristiques d'un "agent" [Cheong, 1996] et dont la réalisation est prévue dans le projet VHF sous la désignation d'un "agent culturel"².

En tenant compte des différents services, le système d'information *desktop* qui est actuellement en cours de réalisation, du côté des partenaires français du projet VHF, recouvre les services (les "applications locales") suivants :

- Recherche d'information encyclopédique (*i. e.* recherche thématique de documents multimédias dans les différentes bases de données du consortium VHF afin de satisfaire une requête donnée) ;
- Recherche de titres d'ouvrages, revues, CDs, vidéos et films, ... saisis dans les différentes bases accessibles ;
- Recherche d'organismes (équipes de recherche, éditeurs, associations, structures artisanales, ...) travaillant dans ou gérant un patrimoine scientifique et/ou culturel ;
- Recherche de compétences "culturelles" ou scientifiques pointues (traducteurs, auteurs, spécialistes de telle ou telle culture, ...).

L'architecture d'un tel système repose sur des bases de données distribuées via Internet et des agents d'informations, c'est-à-dire des programmes spécialisés dans la recherche d'information.

Le fonctionnement général est le suivant :

- l'utilisateur déclenche une ou plusieurs requêtes qui prendront la forme de programmes ("agents") spécialisés,
- les "agents" balayent les documents en interrogeant les bases de données accessibles, collectent les informations recherchées et les communiquent directement à l'utilisateur.

²Pour plus d'informations, cf. le serveur VHF à Naples.

La communication entre ces différents modules impose l'existence d'un module capable de faire comprendre les différents langages et plateformes. Ce module peut être considéré comme un module du type "tableau noir" dans la théorie multi-agents ou comme un module du type "méta-script" dans des modèles distribués tels CORBA ou DCOM.

4. Les services d'informations thématiques

4. 1. Remarques introductives

Dans ce chapitre, nous allons présenter brièvement le service *Dossier Thématique* (cf. 4. 2.) qui propose un ensemble de documents traitant d'une problématique spécifique.

La discussion tiendra compte plus particulièrement des points suivants :

- 1) spécification structurale et fonctionnelle des principaux composants du service en question ;
- 2) simulation de l'utilisation de ces services du point de vue de l'utilisateur ;
- 3) spécification graphique de la maquette des différentes interfaces des services ;
- 4) spécification informatique et implémentation des services.

En revanche, nous ne tiendrons pas compte, ici, faute de place, de la description, pourtant indispensable, sous forme de "cycles de vie" d'un document ou d'un service³.

En problématisant ainsi ce service en question, nous pourrions introduire progressivement la notion du "scénario sémiotique" — notion centrale pour la conception, la spécification et la réalisation d'applications (services et/ou documents) à support électronique (mais aussi "traditionnel").

4. 2. Le service «Dossier Thématique»

4. 2. 1. — Introduction

Le service *Dossier Thématique* est relativement complexe. Il propose trois services spécialisés qui sont les suivants :

- constitution dynamique d'un dossier de documents autour d'une thématique (d'une problématique) choisie ;
- exploration d'une thématique à l'aide d'un scénario d'auteur (*i. e.* d'une structuration de la thématique proposée par un auteur-spécialiste) ;

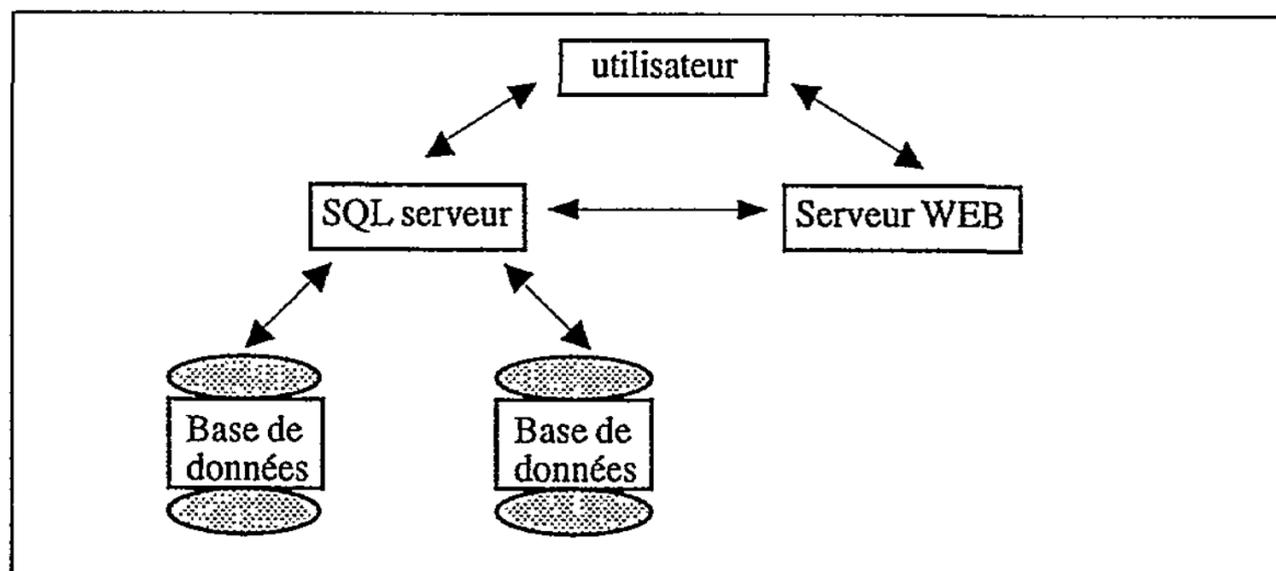
³La notion de "cycle de vie" recouvre les différentes phases d'un document — sa spécification, sa réalisation, sa diffusion, son évaluation et sa maintenance, son archivage et sa destruction. Ils font partie de et sont motivés par le "contrat d'information". Pour une discussion plus détaillée de ces questions, cf. la version électronique de *Semionet* se trouvant sur le site de l'ESCoM.

— production et édition d'un document "personnalisé" (sous forme d'une brochure à support papier ou encore d'un CD-ROM) dans lequel sont réunis les extraits documentaires préalablement consultés par l'utilisateur et dont ce dernier souhaite garder trace.

Nous discuterons ici plus en détail du deuxième type de service qui propose, sous forme d'un scénario d'auteur, une exploration guidée d'un fonds documentaire autour d'une thématique choisie. Toutefois, en ce qui concerne le premier service proposant la constitution dynamique d'un ensemble de documents autour d'une thématique donnée, on peut noter que sa réalisation exige (entre autres) :

- la spécification de la structure conceptuelle et l'implémentation de plusieurs bases de données documentaires qui contiennent un certain type de documents (ex. : base "résumés", base "extraits d'articles", base "extraits d'interviews", bases "références titres" contenues dans le service "éditeurs", etc.) ;
- la spécification de plusieurs index à l'aide desquels les documents stockés peuvent être identifiés et qui servent à la définition des *scripts* (des "programmes") pour la requête ;
- la définition et l'implémentation des *scripts* de requête ;
- la définition et l'implémentation de langages d'interface afin de pouvoir accéder à des documents gérés par d'autres systèmes de bases de données ou encore tournant sur d'autres plates-formes.

Ce service est fondé sur des bases de données en ligne et suppose que tous les documents soient correctement indexés afin de pouvoir appliquer des opérations, de recherche, de jointure, de projection, etc. L'utilisateur spécifie son choix à partir des paramètres représentant les différents types de documents. Le résultat est retourné au format HTML avec des liens actifs ou sous d'autres formats selon un modèle préétabli avec un sommaire, une synthèse, des articles, des publications, des bibliographies, etc. correspondant à la requête envoyée. Les dossiers thématiques générés par le système sont alors prêts pour la consultation, l'édition ou l'impression.



L'implémentation de ce service se fait à partir du modèle des bases de données relationnelles qui, pour chaque type de document, associe des tables. Les *scripts* SQL alors utilisés pour les différents documents sont :

```
CREATE TABLE document
(id_doc          id_doc
doc var_char(20)
type            char(15)
author          char(30)
country         char(20)
date            date
source          text
language        char(10)
region          char(30)
theme           char(40))
```

L'utilisateur choisit d'un certain nombre de paramètres pour générer des documents HTML à partir des BDs :

```
SELECT doc, type, author, country...
FROM document
WHERE critères choisis
```

Au départ le dossier thématique est vide, l'utilisateur le remplit en revoyant ses requêtes et en appelant diverses fonctions. La fonction *INSERT(file, document, part)* insère un document dans un fichier à une position donnée, on alors peut écrire :

```
if document.type=summary
  INSERT(file, document, summary)
if document.type=article
  INSERT(file, document, body)
if document.type=photo
  INSERT(file, document, body)
if document.type=animation
  INSERT(file, document, video)
if document.type=synthesis
  INSERT(file, document, synthesis)
if document.type=bibliography
  INSERT(file, document, bibliography)
```

A la fin de ces opérations, l'utilisateur consulte son "dossier thématique" suivant le format de sauvegarde choisi.

D'autres fonctions sont prévues dans ce service dont la mise à jour des bases de données ; elles sont réservées à un certain type d'utilisateur. La production et l'édition d'un document "personnalisé" sera discuté plus en détail dans le paragraphe 5. consacré au scénario sémiotique.

4. 2. 2. — Le scénario d'auteur

Le scénario d'auteur est un document de type "livre électronique" [Barker, 1991] qui se présente sous la forme d'une application hypermédia

en ligne dont l'originalité repose sur le couplage entre une architecture de base, que l'on peut nommer "classique", et une structure supplémentaire dite "flexible", véritable lieu d'échanges interactifs, ayant comme but de faire évoluer le scénario et de regrouper autour de la thématique développée les spécialistes, les idées, les documents, etc.

Un "livre électronique" ressemble — comme son nom l'indique — à un "livre" à support papier dans le sens où :

— il se construit essentiellement autour d'une "table des matières" (d'un "sommaire") qui propose la structuration thématique d'un domaine d'expertise correspondant au point de vue de l' "auteur" du livre ;

— il offre au lecteur un ou plusieurs chemins de lecture selon lesquels celui-ci acquiert progressivement les informations contenues dans le livre ;

— il propose quelques conseils, aides, explications facilitant la lecture et surtout l'acquisition des informations par le lecteur ;

— il se présente comme un document composé de sections, chapitres, paragraphes et d'une mise en page particulière.

Cela dit, le livre électronique possède, par rapport au livre à support traditionnel, certaines caractéristiques qui lui sont propres, telles que :

a) l'intégration complète de différents médias (texte, images statique et animée, son) ;

b) la multiplicité des formes d'interactivité entre l'utilisateur (le lecteur) du document et le document lui-même, qui ne sont en général pas réalisables dans un document à support traditionnel. Ainsi, il peut jouer à la fois le rôle d'un document didactique à l'aide duquel un utilisateur assimile un ensemble de connaissances sur un domaine, celui du "tuteur" qui contrôle et "guide" l'utilisateur (le lecteur) dans son parcours d'assimilation des connaissances, et celui qui gère et organise l'apport des données extérieures ;

c) la localisation physique de ses données dans des endroits différents (dans différentes bases de fichiers ou de données) — possibilité qui n'est exploitable que très sommairement (sous forme de renvois bibliographique) dans les documents à support traditionnel ;

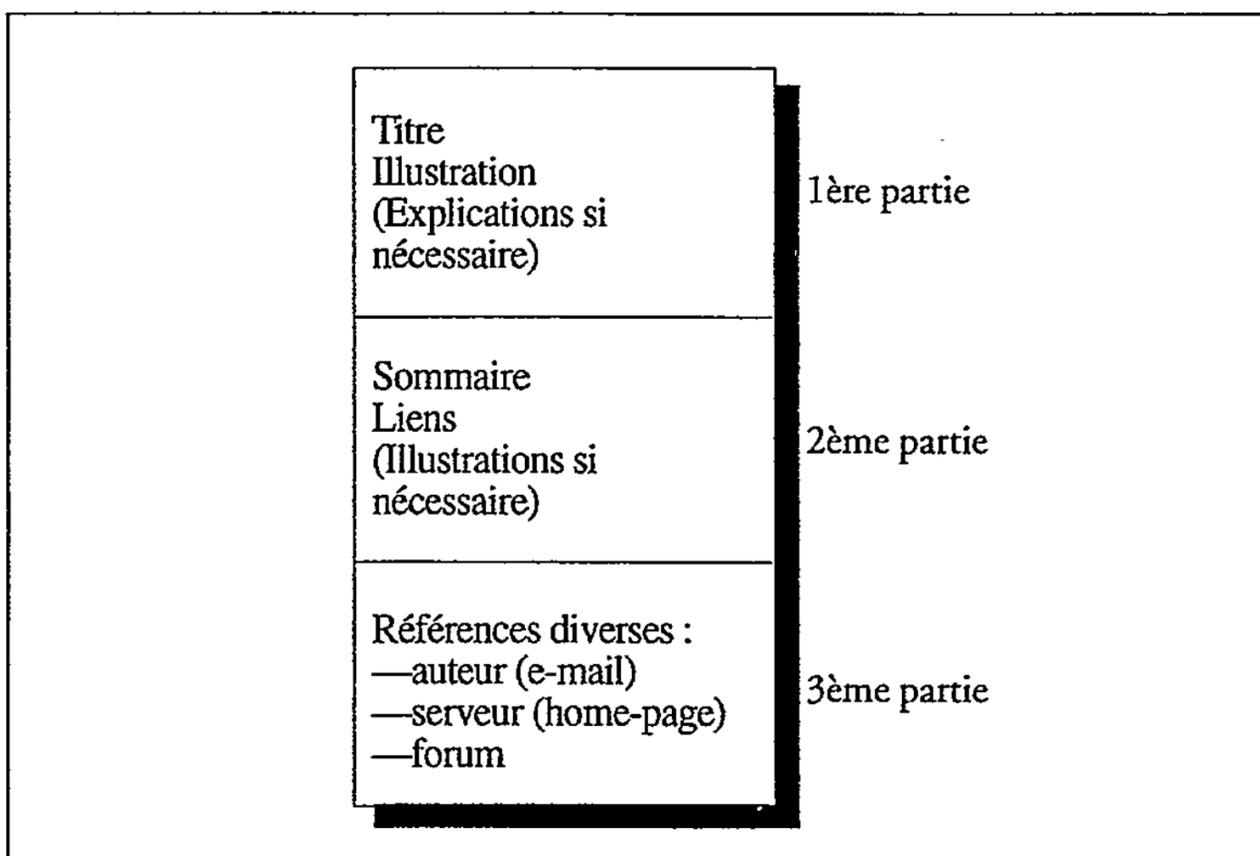
d) une organisation structurale et fonctionnelle suffisamment souple pour être adaptée selon divers contextes d'utilisation et "personnalisée" selon les besoins et les habitudes d'un utilisateur (lecteur) — cette facilité reste relativement limitée pour les documents à support traditionnel.

4. 2. 2. 1. — Le scénario d'auteur «Écriture Chinoise»

Voici l'exemple d'un scénario d'auteur en cours de réalisation au sein du serveur culturel *Orient*. Il s'agit d'un scénario proposant une exploration de différents documents autour de la thématique «Écriture Chinoise».

4. 2. 2. 2. — Organisation et structure de départ du scénario

La structure physique du scénario se présente sous la forme d'une interface simple qui respecte les lignes classiques de composition tout en les adaptant au format électronique :



Suivant cette structure, le sommaire (point de départ du livre électronique) de ce scénario se présente à la manière du tableau ci-après (v. page suivante le tableau du sommaire).

La seconde partie est préalablement établie par l'auteur (= le spécialiste du domaine exploré), mais reste très ouverte dans le sens où le spécialiste :

- organise les informations sources et élabore la structure thématique globale (sommaire, chapitres, sous-chapitres, etc.) ;
- écrit le corps des articles en suivant les standards de conception prédéterminés ;
- propose et crée des liens à l'intérieur du document ;
- compose et remplit les différentes bases de données nécessaires à la gestion des informations et,
- gère l'évolution du scénario (= la partie *flexible*).

Première partie Références introductives (= logo + titre)	Seconde partie Sommaire interactif (= chapitres et sous-chapitres)	Troisième partie Fonctions de navigations et de communications interactives (passerelle vers la structure flexible)
 <p><i>L'écriture chinoise</i></p>	<p><i>Chapitre 1</i> L'origine A. Légendes B. Premières découvertes</p> <p><i>Chapitre 2</i> L'évolution A. Evolution des pictogrammes A1. Les différents styles A2. Synthèse, chronologie illustrée A3. Parties détaillées B. Evolution des supports de l'écriture C. Evolution des politiques liées à l'écriture</p> <p><i>Chapitre 3</i> Les méthodes de classification des caractères A. Shuo wen (540 clefs) B. Méthode des quatre coins C. Kiangxi (214 clefs) D. Simplifiée (189 clefs), etc.</p>	<p>Retour au Home-page Suggestions et commentaires à l'auteur via e-mail. Participation au forum d'échanges. Etc.</p>

⁴Les informations véhiculées seront alors exposées de façon plus formelle voire plus académique ; elles seront établies une fois pour toutes, tout comme dans la construction classique d'un scénario multimédia.

L'application sera alors un produit à visée encyclopédique interactif.

La "restriction d'information" s'établira dans le sens où :

- 1) les liens avec des sites extérieurs et/ou étrangers ne pourront pas être établis,
- 2) les discussions autour de l'application ne pourront se faire en temps quasi réel,
- 3) les évolutions possibles ne pourront s'effectuer autour d'une structure "élastique", il faudra éventuellement qu'elles attendent une modification de version du produit.

La partie flexible est une "sous-structure" ouverte, totalement évolutive, qui permet différents types d'amélioration possibles. C'est un lieu d'interactivité où les informations sont susceptibles d'être modifiées, mises à jour, critiquées, etc. par le biais d'un forum d'échange accessible à tous les utilisateurs qui le souhaitent. Ce forum est composé de quatre "groupes" dont chacun met en exergue une fonction type.

La conception informatique initiale doit donc prendre en charge différents aspects concernant l'architecture classique, où l'interface est préétablie et les informations intégrées, comme l'architecture flexible, où l'interface est préétablie mais où les informations ne sont pas intégrées, ainsi que la mise en place d'un moteur de recherche pour la gestion des informations stockées et/ou récupérées dans différentes bases de données et le maintien du forum d'échange où les "nouvelles" informations sont évaluées et gérées voire récupérées.

Note : Le scénario d'auteur peut également se présenter sous forme de produit hors-ligne de type CD-Rom, CD-I, etc.⁴

4. 2. 2. 3. — Système de consultation du scénario d'auteur

L'utilisateur consulte le document soit en suivant l'ordre donné par le sommaire, point de départ d'exploration du document, soit "dynamiquement" sous sa forme "étendue" d'encyclopédie multimédia évolutive.

Le sommaire permet de visualiser les nœuds composant le document. Chaque chapitre apporte un type d'informations particulières, est composé par n sous-chapitres et constitue donc un mini-document à part entière. Il développe différentes topiques propre à leur thématique. Ces topiques sont exposées dans les sous-chapitres et forment, d'un point de vue sémantique, des nœuds qui peuvent se présenter sous formes de textes, de sons, d'images, de photos, d'icônes, etc.

Voici un schéma simplifié du système de consultation :

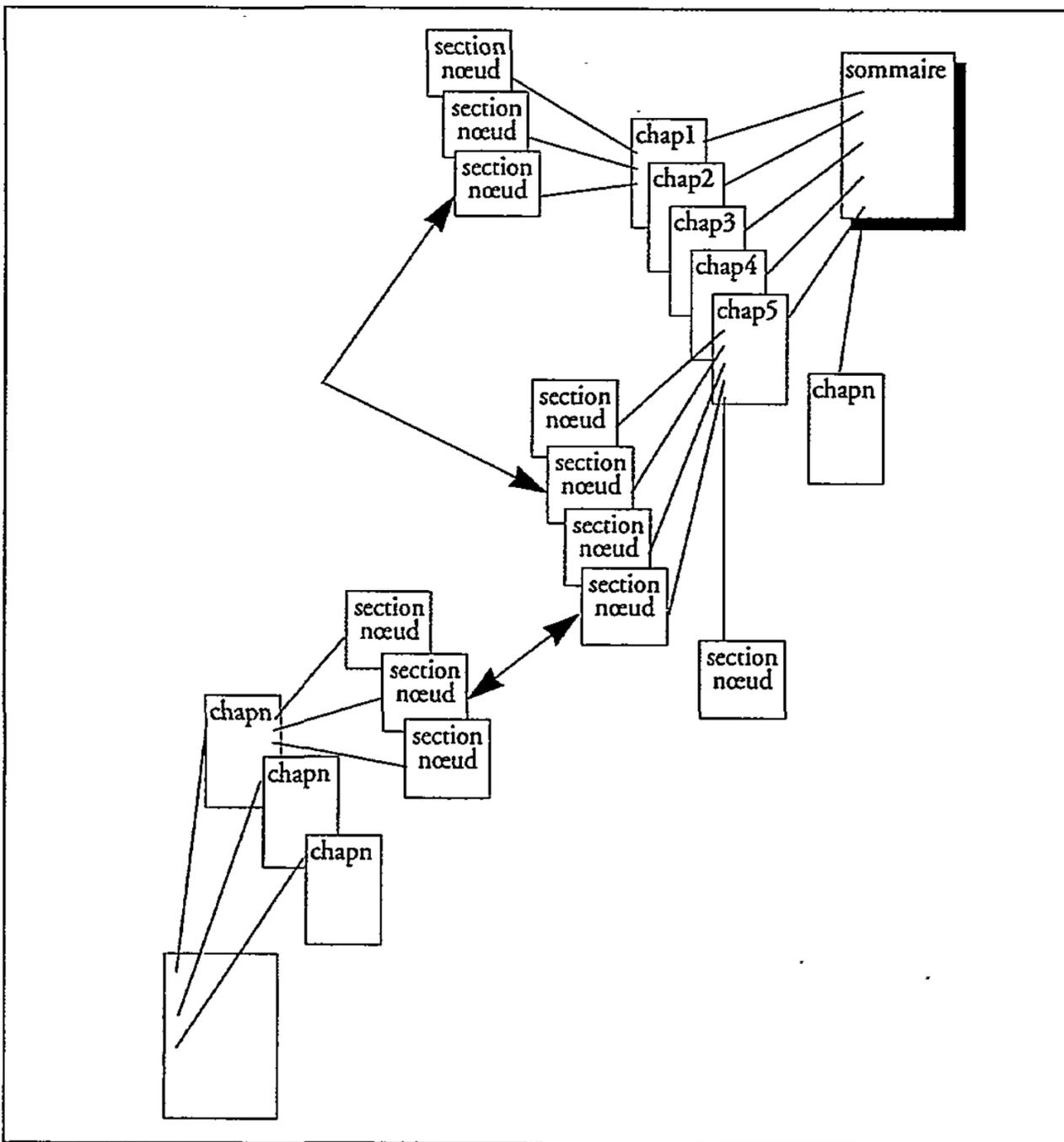


Schéma simplifié du système de consultation

Note : Au plan technique sur le Web, une page ou un écran constitue un nœud. Etant donné que dans le scénario d'auteur chaque chapitre dans sa forme originel constitue une page, il nous est difficile de dire que chaque page constitue une unité de base d'informations, pour cette raison, nous nous référons à un principe d'organisation plus classique de l'hypertexte en considérant chaque chapitre comme un document et chaque sous-chapitres comme un nœud.

Des associations vers d'autres unités informationnelles sont établies à partir de ces nœuds grâce à un ensemble de liens associatifs et référentiels (cf. la structure détaillée ci-avant).

En ce qui concerne la consultation de la partie *flexible*, elle est accessible à chaque chapitre et renvoie vers différentes pages-écrans organisées de façon à visualiser, sous forme de *frames*, les quatre groupes qui la composent. Bien que cette partie constitue l'originalité de ce scénario d'auteur, nous ne la détaillerons pas ici.

4. 2. 2. 4. Organisation générale de la composition physique

Chaque chapitre ou mini-document est donc constitué par :

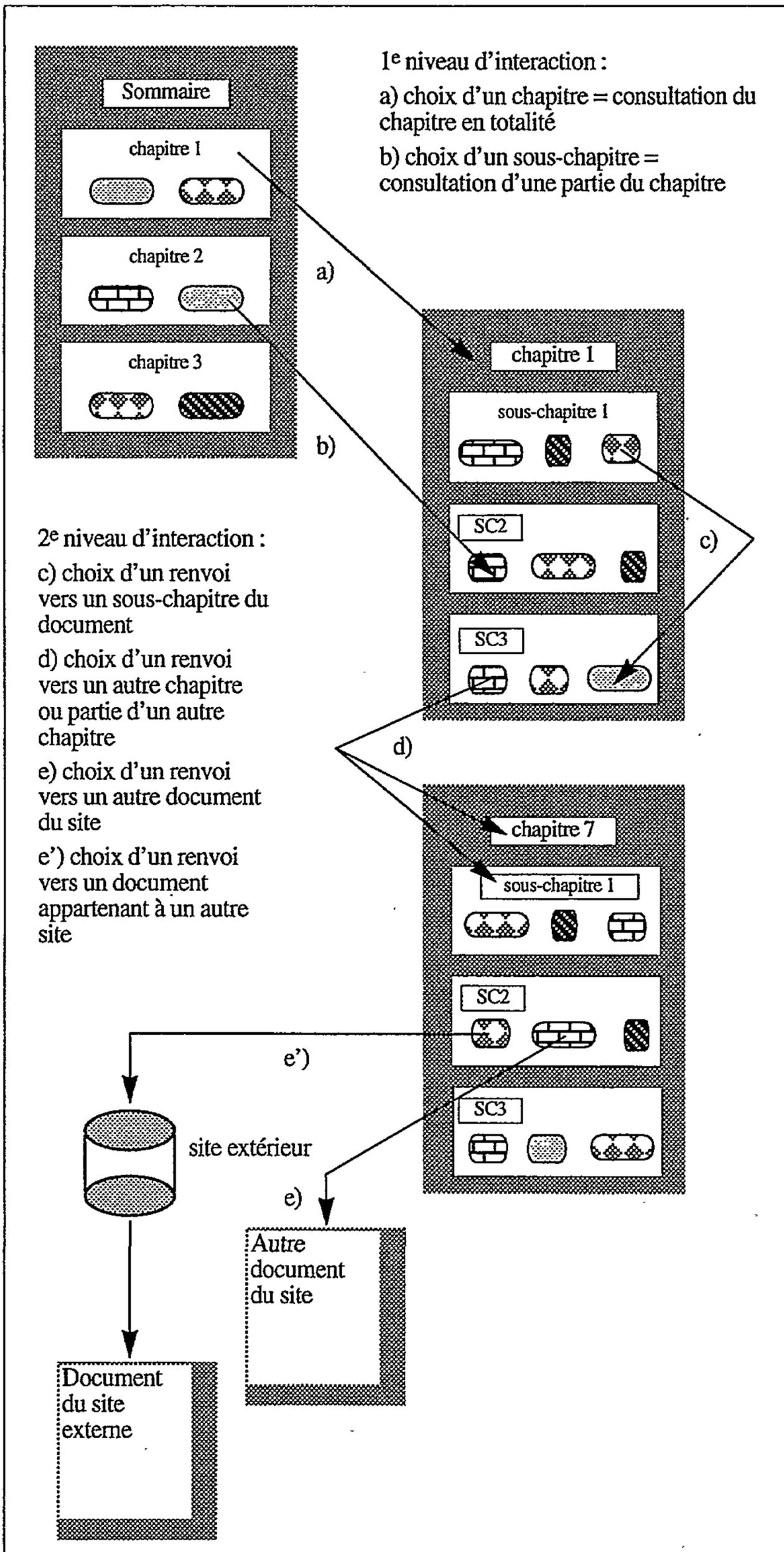
- les différents sous-chapitres ou sections qui forment des nœuds (unité minimale d'informations), ces nœuds peuvent être eux-mêmes formés de différents éléments,
- un ensemble de liens hypertextuels et/ou hypermédias qui permettent une navigation interactive et multimédia à l'intérieur et à l'extérieur du mini-document,
- les différentes parties composant le "forum d'échange".

Les éléments se trouvent physiquement localisés soit "sur place" (directement insérés dans le document), soit "à l'extérieur" dans une base de données qui les cherche et les place dans le document lors de la consultation.

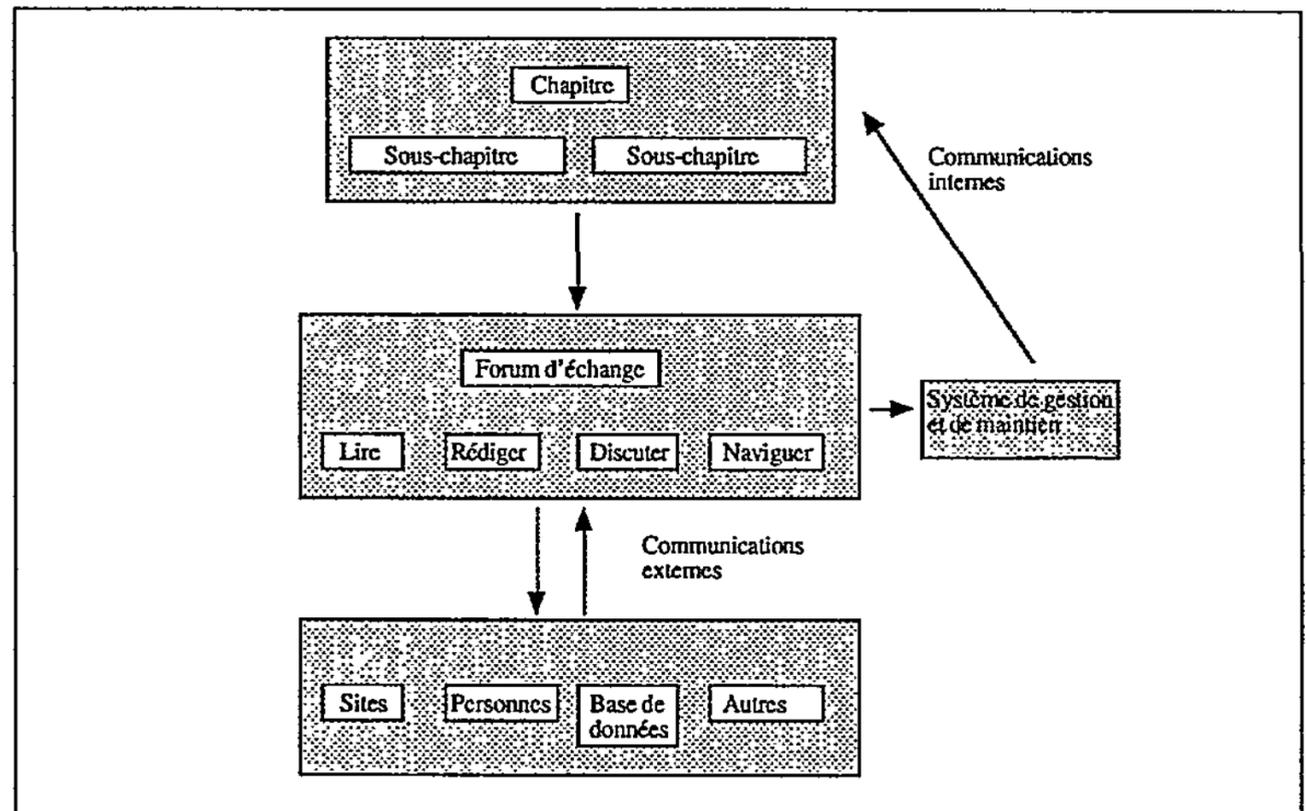
Les liens hypertextuels et/ou hypermédias créent des liaisons entre des documents qui ont ou non les mêmes supports ; ils peuvent être représentés par des mots, des phrases, des adresses électroniques ou encore par des icônes. Ces liens sont dispersés à travers tout le document, ils sont surtout de type associatifs et référentiels.

Des liens organisationnels sont également présents mais ils exercent essentiellement une fonction de navigation à l'intérieur du document : saut de page, retour en arrière, retour au home-page, accès aux différents groupes du forum, etc.

Le graphique ci-contre montre très schématiquement les différents niveaux d'interaction de l'architecture dite *classique*.



Le graphique ci-dessous montre l'architecture *flexible* qui se greffe au schéma précédent.



5. SemioNet : esquisse d'une méthodologie sémiotique

5. 1. Le scénario sémiotique

La notion centrale dans la spécification d'un service est celle du *scénario sémiotique*. Elle a déjà été citée plusieurs fois dans les derniers chapitres lors de la discussion des différents composants constituant un service ou un produit d'informations.

Dans ce chapitre, nous allons introduire plus systématiquement cette notion et discuter son intérêt méthodologique pour la spécification d'un service d'information. Nous insisterons également sur le fait que le scénario sémiotique est un "outil" pouvant être employés dans l'élaboration de différentes tâches.

— Il peut être utilisé pour la spécification de documents et/ou de services à support électronique comme ceux discutés dans cet article mais également pour la spécification de documents à support "traditionnel" (*i. e.* support "papier" et/ou audiovisuel).

— Il peut servir à la définition de la gestion du flux de travail lors de la production d'un service et/ou d'un document concret et donc à la coordination des différentes tâches et activités nécessaires à sa réalisation.

— Il peut enfin constituer un "modèle" de référence pour la description des différents cycles de vie d'un document ou d'un service d'information.

Ce dernier point sera brièvement discuté dans la conclusion consacrée plus particulièrement à l'identification des cycles de vie d'un document dans le contexte du "contrat d'information". Mais pour commencer, essayons d'avoir une idée plus claire de ce qu'est un scénario sémiotique.

Un scénario sémiotique constitue un *modèle* d'un certain type de documents concrets. C'est, en d'autres termes, un *document générique* définissant organisations et fonctions d'une classe de documents qui en sont des *versions concrètes* (ou encore "personnalisées" selon la terminologie utilisée dans les chapitres précédents).

On sait que le scénario, au sens théâtral mais aussi au sens filmique, est canoniquement composé d'un ensemble de scènes (et d'actes) entre lesquelles se tissent, outre une simple relation de succession, différentes relations qui attribuent à chacune des scènes une position particulière dans l'économie globale du scénario. Au sens théâtral ou filmique, une scène est caractérisée, entre autres, par :

— une topique (ou un "sujet") qui peut être identifiée et spécifiée, généralement, autour d'une ou de plusieurs actions, de participants, d'une localisation temporelle dans la structure narrative globale, d'une certaine durée ou encore d'une localisation spatiale ;

— des techniques particulières dont l'objectif principal est de "mettre en scène" la topique, de lui donner une forme particulière (*i. e.* le montage de la topique sélectionnée avec d'autres, sa "prise en charge" subjective, son cadrage et la focalisation sur un thème particulier, son développement linéaire ou non, etc.) ;

— des techniques rhétoriques plus complexes à l'aide desquelles une situation particulière thématifiée est rapportée, commentée, discutée, jugée, louée, etc. par les protagonistes de la scène ;

— des fonctions qu'elle peut remplir dans la structure globale d'un scénario (telles que "introduction", "conclusion", "développement", "digression", "anticipation", ...)

— des modalités de la présentation/expression du contenu par le langage parlé, la voix, la gestualité et la posture, les codes vestimentaires, et ainsi de suite ;

— des instructions scénographiques, à l'adresse du metteur en scène ou encore du lecteur, concernant sa réalisation.

On peut constater également qu'un scénario dramatique peut former un "genre" plus ou moins figé culturellement et historiquement du type "comédie", "western", "policier" ou encore "drame psychologique" servant de "modèle" pour la production de "versions" concrètes (*i. e.* telle

ou telle comédie concrète, tel ou tel policier concret, ...). De même, il est une composition *structurale et fonctionnelle* définissant, comme nous l'avons dit, un certain type de documents concrets, conformément à un objectif d'information donné. Avant de discuter plus explicitement la notion de "composition structurale et fonctionnelle", examinons le cas du document "journal" (au sens de "quotidien d'informations à visée générale").

5. 2. Document générique et versions documentaires

La comparaison d'un corpus constitué de plusieurs numéros concrets d'un quotidien apparus pendant une période X, montrent — exception faite de certains cas plutôt atypiques — des similitudes significatives entre les numéros du corpus qui concernent, par exemple :

- l'organisation et la présentation du "chemin de fer" (*i. e.* l'organisation formelle et physique de l'ensemble des pages constituant un numéro),
- l'organisation et la présentation des pages de couverture (*i. e.* la composition et la présentation de la première page d'un numéro) ainsi que parfois de la deuxième et de la quatrième de couverture,
- le regroupement thématique des informations sous quelques rubriques stables et/ou plus ou moins irrégulières (*i. e.* la "mise en ordre" du jour par un numéro à l'aide d'une sélection et d'une répartition relativement stable et récurrente de certaines grandes topiques),
- la construction et l'emploi de certains genres centraux comme l'éditorial, le "dossier (thématique)" et l' "article de fond", le "point de vue" (d'un expert extérieur) ou encore le "droit de réponse".

Ces similitudes indiquent qu'un numéro particulier constitue la version concrète d'un document générique qui est "le quotidien XX", c'est-à-dire d'un scénario sémiotique dans lequel sont définis : chemin de fer, composition et présentation de chaque page et, plus particulièrement, de la première, deuxième et quatrième de couverture, sélection et répartition des topiques principales, construction des messages journalistiques à l'aide de genre particulier.

La tâche de ce document générique est de définir (*i. e.* de "scénariser") d'une manière plus ou moins contraignante :

- le contenu (*i. e.* la thématique ou encore les topiques) pouvant être sélectionné et développé par une version concrète ;
- les genres utilisés — aussi bien généraux (description, narration, débat,...) que spécifiques, voire très spécifiques tels que l'éditorial,

l'article de fond, le droit de réponse, ... ;

— le ou les langages (conventions langagières) à respecter lors de la rédaction d'un article, d'un numéro ;

— les formes de présentation (les conventions) typographiques et chromatiques ;

— les illustrations graphiques (schémas, dessins, caricatures, ...) et photographiques ;

— l'organisation formelle et physique (*i. e.* mise en page et chemin de fer) des différents numéros ;

— l'organisation formelle et physique des différents types de pages constituant un numéro.

Nous noterons que, contrairement à une compréhension fautive de la notion "générique", le document générique n'est pas forcément un document abstrait ou général, il est tout simplement *le modèle* ou encore *un modèle de référence* suivant lequel un nouveau document — une version du document générique — est produit. Le document générique est, autrement dit, une *norme* (un *standard*) qui autorise la production d'une certaine variété de versions documentaires.

La norme elle-même peut être très spécifique et ne permettre que des versions légèrement différentes les unes des autres par rapport au document générique lui-même. Le cas limite est qu'une version concrète du document générique ne soit que sa *parfaite copie*.

Le document générique peut, bien sûr, posséder une relativement grande généralité spécifiant, d'une manière très sommaire, contenu, présentation et organisation d'une version donnée mais qui a l'avantage de proposer un schéma global pour une très grande variété de versions documentaires dont la ressemblance n'est pas évidente.

De même, le document générique constituant la norme pour une certaine variété de versions documentaires, peut — d'une manière parfaitement similaire à tout document — posséder une *structure de complexité* plus ou moins importante : il peut être constitué par des composants qui l'organisent localement et qui, de leur côté, peuvent être, indépendamment des autres composants, plus ou moins spécifiés, restreints à quelques cas typiques ou, au contraire, rester relativement schématiques et laisser ainsi une grande liberté aux versions pour lesquelles ils constituent un modèle partiel.

5. 3. Document générique et définition d'un type de documents en SGML

Le document générique correspond, *grosso modo*, à ce que l'on appelle une "définition d'un type de documents" (*Document Type*

Définition, DTD) dans le standard SGML (*Standardized General Mark up Language*) — standard utilisé dans la gestion électronique de documents pour structurer à l'aide d'un ensemble de balises (*tags*, en anglais) le document au niveau de la présentation des codes d'expression, de l'organisation en chapitres, paragraphes, titres, etc., ainsi qu'au niveau du contenu (de la topique) du document [Travis, Waldt, 1995].

Une telle structuration doit rendre plus ou moins indépendante l'organisation du document de la plate-forme logiciel utilisée et permettre aussi bien l'utilisation d'un document sur différentes plates-formes par des environnements logiciels différents, que la réutilisation partielle d'un document (sous forme de composants documentaires) lors de la production d'un nouveau document, la maintenance locale d'un document déjà existant ou encore l'édition d'une version "personnalisée" d'un document existant.

Un DTD, dans sa forme la plus simple, correspond à un sommaire qui détaille, sur un niveau de généralité, aussi bien le contenu des différents chapitres ou paragraphes constituant un type générique de documents que la position (hiérarchique et linéaire) de chaque chapitre par rapport aux autres. Illustrons notre propos par le DTD d'un vade-mecum pour touristes qui est en train d'être développé dans le cadre du projet Européen VHF.

Ce vade-mecum doit constituer une petite brochure ou un CD-ROM dans lequel le touriste trouve un ensemble d'informations utiles pour son voyage et/ou son séjour dans un pays. Il représente une partie du service en ligne appelé «Valise pour Touriste» dans lequel est proposé la consultation d'informations pratiques concernant la préparation d'un voyage et d'un séjour dans un pays ainsi que des offres de voyage de quelques opérateurs touristiques. Une fois que l'ensemble des informations choisies par l'utilisateur est consulté, ce dernier peut, s'il le souhaite, procéder à une récapitulation des documents visualisés (à une sorte d'évaluation de son parcours) et indiquer les documents qu'il aimerait voir assemblés — au sein d'une petite brochure ou d'un CD-ROM simple — dans le vade-mecum en question.

Les informations sauvegardées par l'utilisateur vont être *assemblées* dans une structure, très simple, de document prédéfini (la fameuse DTD) qui caractérise l'organisation générale du vade-mecum. Chaque composant de cette structure globale est, à son tour, spécifié à la fois d'un point de vue de la maquette, du contenu et de l'expression. En ce sens le chapitre «Offres de Voyage» comporte des composants tels que "deuxième titre" (*i. e.* titre de l'offre de voyage), "paragraphe-chapeau" (*i. e.* paragraphe donnant un résumé d'un voyage ou séjour), "paragraphe-séjour" (*i. e.* texte développant le séjour proposé dans offre de voyage), "paragraphe-carte" (*i. e.* zone réservée à la carte illustrant l'itinéraire), etc.

La production d'un vade-mecum (d'une brochure concrète) suit *grosso modo* les étapes suivantes : (1) recherche des informations

pertinentes dans les bases de données (internes, pour le moment) ; (2) identification des données documentaires contenant ces informations et transfert dans le module de la production du vade-mecum ; (3) sélection du scénario «Offres de voyage» ; (4) assemblage des données documentaires selon la structure générique propre au scénario ; (5) production des pages contenant les informations pertinentes. Ce processus ne se déroule pas d'une manière linéaire et déterministe mais l'espace nous manque ici pour en discuter les détails.

5. 4. Le scénario comme composition structurale et fonctionnelle

Comme nous l'avons dit plus haut, le scénario sémiotique est une composition structurale et fonctionnelle. En tant que *composition structurale*, il spécifie un type de document (comme justement le quotidien XX) par rapport aux principaux plans qui caractérise, par définition, tout type de documents :

— le *plan du contenu* a comme tâche la saisie des informations thématiques et développées par un document, voire par un corpus de documents ;

— le *plan des codes et médias de présentation* du contenu a comme tâche la saisie des modalités et des stratégies d'expression d'une information thématique et développée par un document, un corpus de documents ;

— le *plan de l'organisation formelle et physique* du document a comme tâche la saisie des composants documentaires (cf. *infra*) qui constituent et structurent un document ;

— le *plan du support* (du matériel) a comme tâche la saisie — comme le nom l'indique — du médium physique à l'aide duquel un document se réalise et se présente en tant que produit concret.

Il faut noter ici que la distinction entre ces quatre plans est plutôt motivée par des raisons pratiques, voire technologiques liées à la production et l'évaluation d'un document. Elle n'est, cependant, pas contradictoire avec la fameuse distinction structuraliste plan du contenu ("signifié") et plan de l'expression ("signifiant") dans la mesure où le premier plan cité ci-dessus correspond *grosso modo* au plan du contenu (au signifié), le deuxième et troisième plans correspondent au plan de l'expression et le plan du support physique correspond au plan de la substance de l'expression (dans la terminologie de Hjelmslev [1968 (1943)]).

D'un point de vue *fonctionnel*, le scénario sémiotique définit un type de documents par rapports aux grandes *tâches* qu'ils sont censés remplir

selon un objectif d'information. Parmi les fonctions principales, on compte :

— la *fonction textuelle* qui spécifie le propos d'un document ou d'un type de document et donc, *a fortiori*, la sélection et la composition (le montage) des topiques ou thématiques dans lesquelles puisent l'auteur afin de réaliser son propos ;

— la *fonction métatextuelle* qui spécifie les différentes formes d'interactions entre le destinataire d'un document et l'auteur de celui-ci en vue de réaliser avec succès le transfert d'une information. Dans cette classe de fonctions, on range, en outre, l'aide et l'explication, la gestion de l'accès et de la lecture d'un document, les libertés offertes au lecteur/utilisateur de modifier un document ;

— la *fonction hypertextuelle* qui spécifie les différents moyens et procédures pour compléter les informations données dans un document avec des informations se trouvant dans d'autres documents. Dans cette classe de fonctions, on compte, en outre, la bibliographie, l'iconographie, la navigation dans un hyperdocument (*i. e.* un document qui se réalise par l'assemblage dynamique de composants documentaires qui sont physiquement différents) ;

— la *fonction péritextuelle* qui spécifie la "clôture" d'un document par rapport à son "environnement" et donc son identité. Dans cette classe de fonctions, on range, entre autres, les indications concernant l'auteur (les auteurs) d'un document (des composants documentaires constituant un hyperdocument), les précisions concernant le titre ou le générique (dans le cas de documents audiovisuels), des précisions concernant l'éditeur, l'imprimeur et la date de parution, des précisions en ce qui concerne le type d'édition, et ainsi de suite).

Comme nous le verrons encore, chaque composant d'un scénario sémiotique doit être décrit aussi bien d'un point de vue structural que fonctionnel. En d'autres termes, un composant sémiotique donné est spécifié par rapport à l'un ou l'autre des quatre plans constitutifs ou par rapport à plusieurs plans à la fois ; il est également spécifié quant à son rôle, et quant à la fonction qu'il est censé remplir dans le scénario.

5. 5. Scénario global et scénarios locaux

Dans le cas de documents complexes, on peut aisément distinguer entre plusieurs scénarios qui composent ensemble la structure générique des versions concrètes de tels documents. Pour reprendre le cas du quotidien, il est assez simple de pouvoir distinguer :

— des *scénarios locaux* définissant par exemple la page de couverture, telle ou telle rubrique, tel ou tel dossier thématique, la page des petites annonces, la page réservée aux lecteurs, etc. ;

— et un *scénario global* (voir plusieurs scénarios globaux) définissant la place et la fonction des différents scénarios locaux dans la structure globale.

Cette organisation très *modulaire* du scénario sémiotique permet de tenir compte (*i. e.* de décrire, d'expliquer mais aussi de planifier) des variations des versions concrètes d'un document générique — exigence qui est particulièrement importante pour la gestion de documents électroniques comme, par exemple, d'un journal électronique ou encore d'un manuel électronique dont l'un des objectifs principaux est l'adaptation à des besoins, habitudes ou encore désirs plus ou moins spécifiques d'un utilisateur.

La spécification structurale et fonctionnelle des différents types de sites spécialisés, hébergés par le serveur *Orient*, tient justement compte de l'adaptation d'un type de site donné comme, par exemple, du type «site de recherche» ou encore du type «site d'éditeurs» aux particularités de telle ou telle équipe de recherche, de tel ou tel éditeur.

Cela signifie que chaque type de sites est structurellement et fonctionnellement spécifié par un ensemble de composants (de "modules") qui constituent un site et y occupe un rôle particulier. Ces modules constituent en quelque sorte une *bibliothèque* de composants préétablis qui sont sélectionnés, partiellement modifiés et assemblés d'une part selon un scénario donné et d'autre part selon les particularités de tel ou tel éditeur, de telle ou telle équipe de recherche.

De même, une application à visée encyclopédique et ou à thème comme le *Dossier Thématique* présenté dans le paragraphe 4., est la réalisation d'un scénario sémiotique, c'est-à-dire d'un type générique d'applications à l'aide duquel on peut facilement produire et réaliser d'autres services et documents à visée encyclopédique et à thème qui seraient plus ou moins proches de ces dernières. Cette facilité est au cœur de la production et réalisation de sites *Web* à thème.

5. 6. Problématisation et maintenance d'un scénario sémiotique

En tenant compte de la distinction capitale entre type générique de documents et/ou services et versions (personnalisées, concrètes, ...) d'un tel type générique, deux axes de recherche et de développement complémentaires se dégagent :

— le premier vise à la fois la problématisation et la maintenance du scénario sémiotique,

— le second tient compte des procédures de production et de réalisation d'une version particulière d'un scénario sous forme d'un document ou service concret.

Nous reviendrons, plus loin, sur la question de la réalisation d'une version particulière à partir d'un scénario sémiotique. Examinons rapidement le premier axe de recherche et de développement.

La problématisation et la maintenance du scénario sémiotique concernent, en premier lieu, l'examen critique des différents composants — des modules — qui constituent un type générique de documents et services. Si nous reprenons l'exemple du scénario d'auteur présenté dans le paragraphe 4., l'examen critique concernera les points suivants :

— les "rubriques" thématiques constituant le document en question sont-elles les seules possibles ou faut-il en spécifier d'autres (non prises en considération) ? et si oui, comment s'insèrent-elles dans les rubriques déjà existantes ?

— Les chemins ou parcours établis entre les différentes rubriques, sont-ils les seuls possibles dans le scénario ? ou, au contraire, peut-on (doit-on) en définir d'autres ?

— Les codes et langages d'expression utilisés dans le document réalisé à la fois pour la communication et la consultation des informations doivent-ils être considérés comme les seules formes d'expression ou faut-il réserver, dans le scénario sémiotique, des formes d'expression alternatives pouvant satisfaire davantage tel public ou tel autre ?

— De même, quant à l'organisation formelle et physique des pages-écrans, doit-elle rester toujours identique à celle choisie pour le document réalisé ou ne peut-on (ne doit-on) pas prévoir des standards de pages-écrans différents afin de répondre mieux aux habitudes ou attentes de lecture d'un public donné ?

— Enfin, ne faut-il pas identifier et spécifier de nouveaux composants qui ne sont pas intégrés dans le document afin de pouvoir produire des versions du scénario d'auteur qui tiendraient davantage compte d'une situation d'apprentissage, d'un débat autour d'un thème précis ou encore de la production collective d'un document nouveau traitant d'un aspect particulier de cette thématique ?

Il s'agit ici de questions typiques qui concernent en premier lieu la problématisation d'un scénario et sa mise à jour. Mais "problématisation" veut aussi dire, développement de théories descriptives et formelles plus sophistiquées sous-tendant la construction et la maintenance de scénarios

sémiotiques. De tels efforts théoriques concernent la description et la modélisation thématiques (topiques) nécessaires pour le développement et la maintenance de thesauri et d' "agents intelligents" de recherche d'information sur le *Web*, la description de plans thématiques sous-tendant la mise en séquence et la construction (linéaire) de documents, la description de codes d'expression d'un contenu et des contraintes intersémiotiques entre différents codes, ou encore la description des maquettes de documents et de leurs composants.

Pour mieux apprécier ces questions et les enjeux aussi bien théoriques que pratiques, il nous faut développer les deux notions de "composant documentaire" et d' "unité scénarielle".

5. 7. Composants documentaires et unités scénarielles

Reprenons notre exemple du quotidien XX introduit auparavant, et précisons que pour caractériser les composants du scénario en question nous avons eu recours :

- d'une part à trois des quatre plans constitutifs d'un document (nous avons laissé de côté le plan "support") ;
- d'autre part à des "unités nommées" tels que "dossier", "éditorial", "enquête", "article de fond", "titraillé", "chapeau", "récit" ou encore "évaluation".

Nous appelons les unités nommées "composants documentaires" ou encore, dans le cas de documents multimédias, "unités scénarielles".

Un composant documentaire se distingue par le fait qu'il possède une structure, une organisation relativement figée et facilement reconnaissable par un lecteur "moyen" — il se distingue par ce que Eco a appelé un *code* [Eco, 1972].

Des composants tels qu'un éditorial, un dossier thématique, une enquête, un article de fond sont donc définis par un *scénario assez typique* qui, certes, peut varier d'une version concrète à une autre, mais qui garde cependant une certaine stabilité faisant en sorte que chaque version concrète est interprétée comme appartenant au type "éditorial", au type "dossier thématique", etc.

Le scénario d'un composant documentaire peut être spécifié davantage soit sur l'un ou l'autre des plans constitutifs d'un document, soit sur plusieurs à la fois. Par exemple :

- les composants du type "éditorial", "article de fond", "brève", "sommaire", ... sont des composants spécifiés surtout sur le plan thématique ;

— les composants du type “première” et “deuxième” couverture, encadré, paragraphe, sont surtout spécifiés au niveau de l’ “organisation formelle et physique” (*i. e.* mise en page) et de la présentation graphique ;

— les composants du type “schéma”, “police”, “bordure”, ... sont surtout spécifiés au niveau de la présentation typographique et chromatique.

Les composants documentaires introduits ci-dessus montrent que l’on peut distinguer entre différents types de composants documentaires tels que :

— des types motivés surtout d’un point de vue topique (cf. le cas du composant “rubrique”) ;

— des types motivés du point de vue de la composition thématique (cf. le cas des composants comme “dépêche”, “filet”, “article de fond”, etc.) ;

— des types motivés aussi bien du point de vue de la présentation que de l’organisation formelle (cf. le cas des composants “titre”, “paragraphe”, “encadré”, ...)

— des types motivés du point de vue de la maquette, la mise en page et la position plus ou moins invariable dans le chemin de fer (cf. le cas des composants “page de couverture, quatrième de couverture”, etc.).

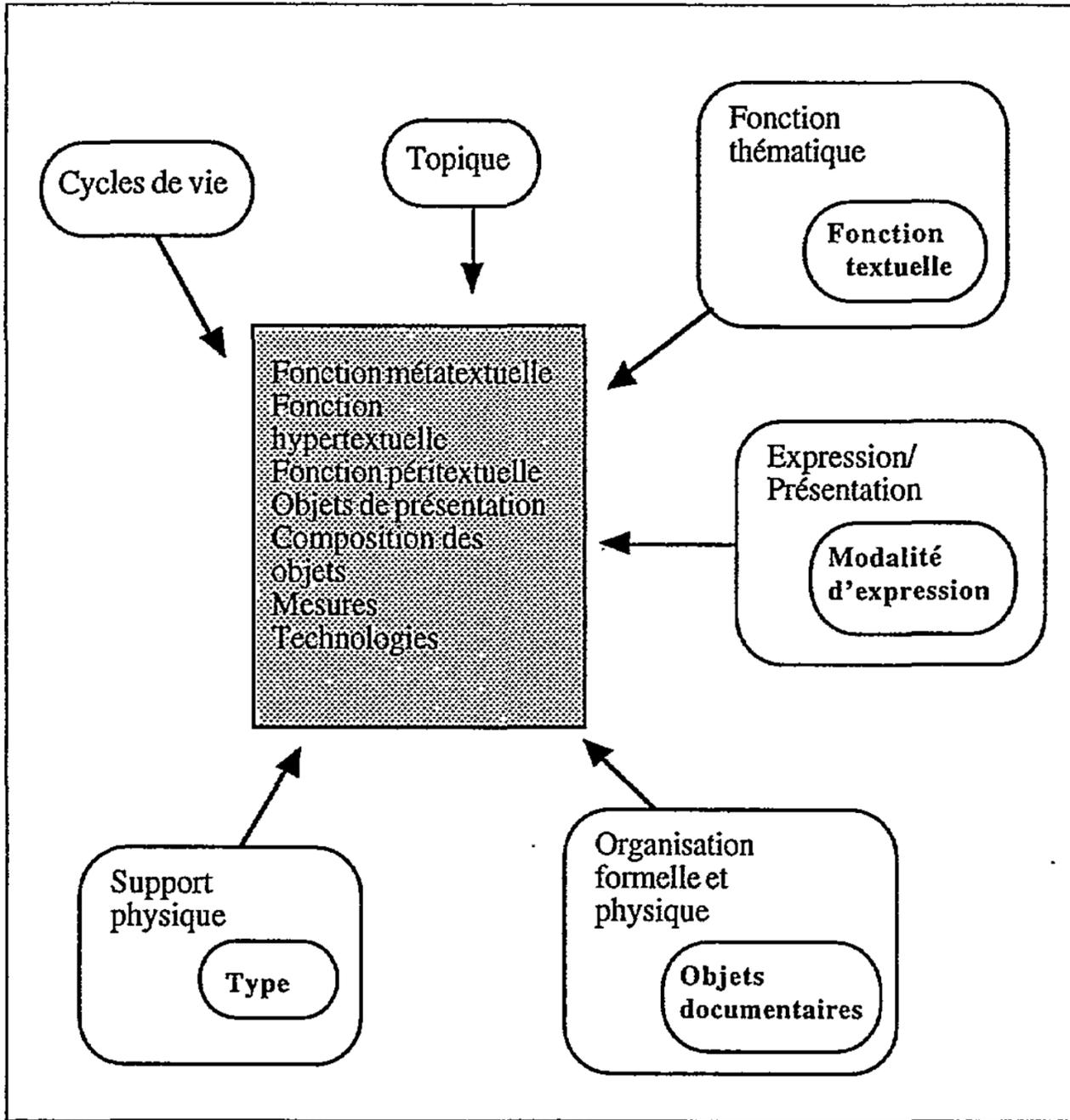
Chaque type accomplit selon sa spécificité, des *tâches particulières* dans la production, la réalisation du document “journal”. Etant donné leur spécificité :

— les composants “à dominante topique” (*i. e.* motivés surtout d’un point de vue de la topique), organisent et guident la production du document au niveau de l’identification, de la sélection et de la classification des informations à traiter ;

— les composants “à dominante composition thématique” organise et guide le traitement des informations thématiques selon tel ou tel genre ;

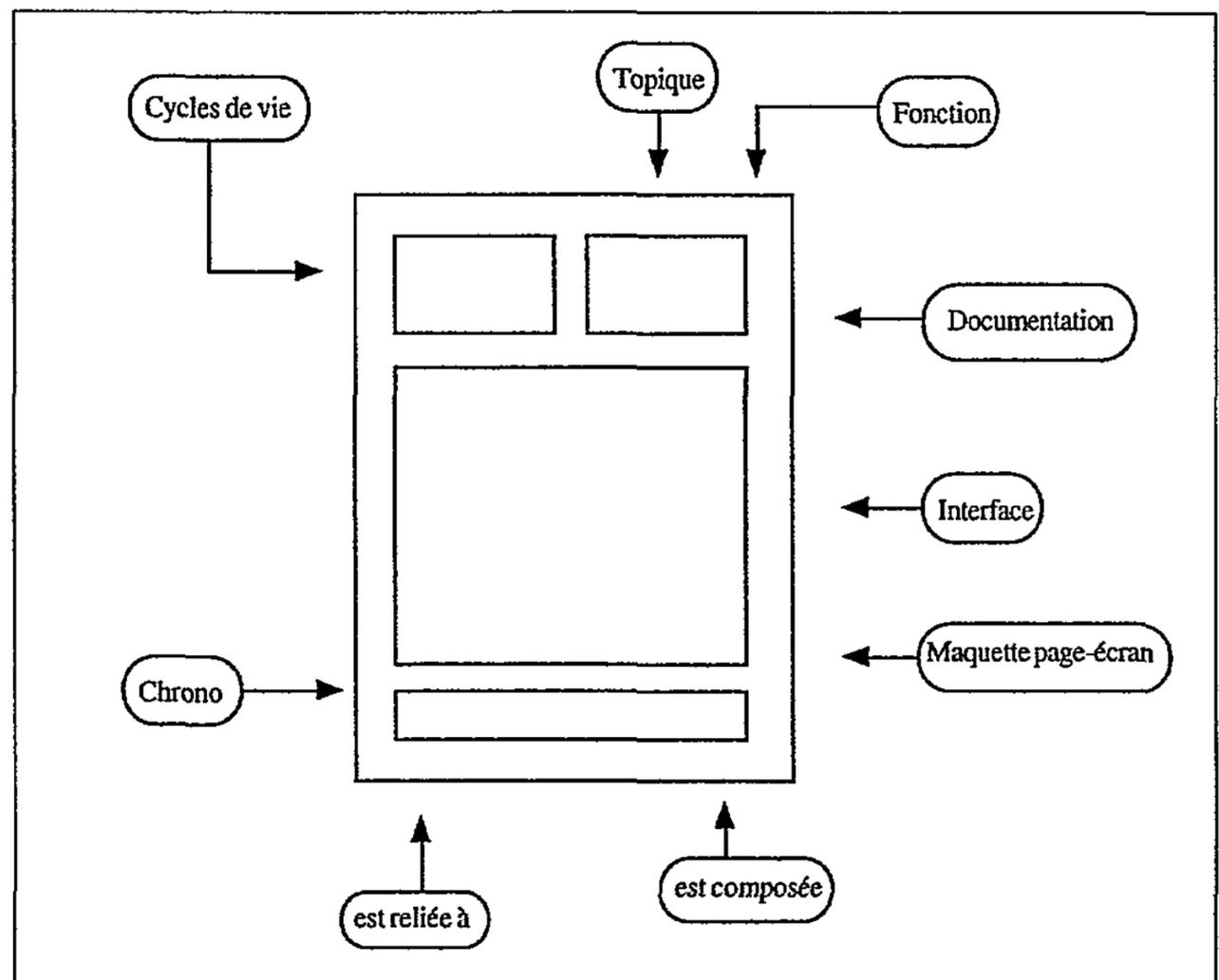
— les composants “à dominante présentation et maquette” organise et guide le choix des médias d’expression et la mise en page des informations thématiques et traitées par tel ou tel genre.

Le schéma ci-contre explicite les différents paramètres à prendre en compte lors de la spécification ou de l’évaluation d’un composant documentaire.



Dans ce schéma, sont identifiés également les différents critères à prendre en compte lors de la spécification d'un composant documentaire. Spécification voulant dire ici, conception d'un scénario qui sous-tend la production-réalisation de versions (concrètes, "personnalisées").

Comme nous l'avons déjà mentionné, l'unité scénarielle est le composant documentaire des applications audiovisuelles et/ou multimédias (*i. e.* à support électronique) — applications dont la réalisation doit prendre en considération la dimension temporelle (à propos de l'unité scénarielle, cf. aussi [Stockinger, 1992, 1993]). Le schéma ci-après représente les critères de spécification de l'unité scénarielle qui sont très similaires à ceux utilisés pour la description du composant documentaire (en d'autres termes, l'unité scénarielle est un type de composant documentaire particulier).



Une unité scénarielle (US) est un objet paramétré composé des parties suivantes :

- partie réservée aux informations signalétiques telles que : titre de l'US, mot(s) clé, numéro d'ordre, numéro de version, auteur/concepteur et date ;
- partie réservée à l'identification et à la définition de la topique traitée dans l'US ;
- partie réservée à l'identification fonctionnelle de l'US (*i. e.* quel est le rôle censé jouer par l'unité dans l'application ?) ;
- partie réservée à l'identification des unités scénarielles pouvant être atteintes à partir d'une unité départ (*i. e.* avec quelles US, l'US est "mise en réseau" ? quels sont les chemins empruntés par l'utilisateur à partir de l'US en question ?) ;
- partie réservée à la "documentation" de la topique traitée dans l'US en question (*i. e.* les textes, graphiques (schémas, diagrammes, tableaux, ...), dessins, photos, vidéos, discours, musiques, bruits, ... utilisés pour développer et documenter une topique choisie) ;
- partie réservée à la spécification des interfaces (iconiques, textuelles, sonores, ...) ;
- partie réservée à la spécification de la maquette des pages-écrans (structure et organisation de chaque page-écran et succession des pages-écrans) ;

- partie réservée à l'identification d'US qui composent l'unité en question et qui possèdent une structure autonome par rapport à celle-ci ;
- partie réservée au chronométrage de l'US et à la synchronisation des différentes parties constituant l'US en question ;
- partie réservée aux effets de transition déterminant les passages entre les différentes parties constituant une US ainsi que ceux entre une US et d'autres US.

Les modules et autres composants des différents services du serveur culturel *Orient* sont spécifiés et enregistrés de cette façon et constituent ainsi une bibliothèque de composants destinée à la production, la réalisation ou encore la maintenance d'un (nouveau) service ou document.

5. 8. La production-réalisation d'une version documentaire

Le scénario sémiotique peut être comparé à un document actualisé, c'est-à-dire à un document qui à la fois précède et circonscrit la production d'un ensemble ouvert de documents concrets qui en constituent autant d'exemples.

Grosso modo, un document concret (réalisé sur un support physique particulier) peut être considéré comme le résultat d'un ensemble de décisions portant, entre autres, sur :

- la sélection de composants documentaires prédéfinis et stockés, par exemple, dans une ou plusieurs bases de données ;
- l'assemblage (le montage) des composants conformément à des objectifs d'informations à réaliser et à l'organisation structurale et fonctionnelle du scénario sémiotique. Cette opération étant similaire à celle de la création d'une maquette (pour des documents écrits) et d'un *storyboard* (pour des documents audiovisuels) ;
- les mises à jour ou modifications propres à chaque composant sélectionné. Les mises à jour ou modifications concernent le contenu d'un composant sélectionné (exemple : montage des thèmes d'une topique à l'aide de thèmes d'une autre topique, amplification de certains thèmes, séquentialisation d'une topique selon des "maximes" ou "règles" particulières, et ainsi de suite), les modalités de présentation et d'expression (exemple : modification du composant chromatique ou figuratif d'une image, modifications typographiques, ...) ainsi que l'organisation formelle et physique d'un composant (exemple : redéfinition de cadres, de la taille d'une image, du nombre de colonnes, de la synchronisation entre l'image et la voix, etc.).

Notons ici seulement que cette "vision" est assez similaire au processus de création-production d'applications multimédias simples et qu'il sous-tend l'utilisation d'un générateur d'applications tel que *Director* de *Macromedia*.

En effet, "l'environnement de travail" proposé par *Director* est constitué par le *cast*, le *score* et le *stage*.

— Le *cast* est la partie de l'environnement de travail dans laquelle sont stockés les différents composants documentaires (allant des composants les plus simples comme les bordures, les fonds, ... jusqu'au composants les plus complexes tels que des textes, des images, des vidéos, ...) devant servir lors de la production d'un nouveau document selon un scénario sémiotique préalablement spécifié.

— Le *score* constitue le *storyboard* dans lequel sont assemblés les composants documentaires sélectionnés et "stockés" dans le *cast*. L'assemblage se fait selon la "logique" spécifiée préalablement dans le scénario sémiotique.

— Le *stage* (la "scène") constitue la partie dans laquelle on peut procéder aux modifications indiquées ci-dessus afin d'adapter les composants documentaires sélectionnés aux contraintes du scénario sémiotique préalablement spécifié.

Enfin, pour créer des applications documentaires plus sophistiquées répondant aux critères de l'interactivité et de la personnalisation, *Director* propose un langage de programmation (*Lingo*) à l'aide duquel on peut réaliser des *scripts* qui — au moins jusqu'à un certain niveau de complexité — peuvent tenir compte de telles exigences.

6. Conclusion

Bien que nous ayons mentionné la question des *cycles de vie* d'un document, d'une application ou encore d'un service particulier lors de la présentation du projet Européen VHF et des services constituant le serveur *Orient*, nous en avons fait largement abstraction.

Nous ne pouvons, faute de place, détailler cette question. Nous nous contentons donc des quelques remarques suivantes — remarques qui ont comme but de contextualiser, dans un ensemble d'autres activités et méthodes, le scénario sémiotique comme méthodologie et outil de spécification (conceptuelle) de documents et/ou de services.

Un document — comme un service — doit répondre à un objectif d'information. L'objectif d'information est défini dans ce que qu'on peut appeler le *contrat d'information*. Le contrat d'information est une

construction théorique qui essaie de tenir compte des enjeux et des phases principales caractérisant l'émergence d'un manque ou d'un besoin d'information, des modalités de sa réalisation, de sa réalisation proprement dite et enfin de l'évaluation et sanction de la réalisation.

Le contrat d'information peut être décomposé en plusieurs phases principales dont chacune remplit une fonction particulière dans la production d'un document, c'est-à-dire d'un service ou d'un produit d'information. Ces différentes phases conditionnent ce que l'on appelle la description d'un document ou d'un service d'information à travers ses *cycles de vie* — *i. e.* de sa "naissance" jusqu'à sa "mort".

La notion de "cycle de vie" possède des acceptions différentes : parfois elle signifie la "vie" d'un document depuis sa réalisation jusqu'à sa fin (jusqu'à son archivage ou sa destruction), parfois elle inclut le processus de la création d'un document à partir du simple constat d'un besoin d'information. En privilégiant l'acceptation plus large de "cycles de vie" d'un document (ou d'un service), nous pouvons distinguer plusieurs phases, activités et tâches caractérisant la "vie" d'un document (ou d'un service) que l'on peut regrouper autour des pôles suivants :

- (I) «Analyse des besoins en informations et de la demande en produits documentaires et/ou services d'information»,
- (II) «Analyse de l'existant et de l'offre en services et produits d'information» afin de pouvoir déterminer si les besoins analysés en (I) peuvent être satisfait par des produits ou services existants,
- (III) «Spécification (conceptuelle, technique, ...) d'un produit ou service d'information (partiellement) nouveau» devant répondre aux besoins analysés en (I),
- (IV) «Spécification "contractuelle" pour la réalisation, la validation, l'intégration, la maintenance et l'exploitation du nouveau service et/ou produit d'information»,
- (V) «Réalisation du service ou du produit d'information» selon (III) et (IV)
- (VI) «Test, validation, correction» du service ou du produit d'information»,
- (VII) «Maintenance du service ou du produit d'information (y compris l'archivage et la suppression du produit ou du service d'information)»,
- (VIII) «Exploitation du nouveau service ou produit d'information».

La place principale du scénario sémiotique dans l'ensemble de ces activités et tâches se trouve dans le pôle (III), c'est-à-dire dans le pôle qui

comprend les activités et tâches visant la *spécification conceptuelle et technique* d'un (nouveau) document ou d'un (nouveau) service. En tant que *modèle de référence*, le scénario, lors de la réalisation, de l'évaluation, de la maintenance et de l'exploitation d'un produit ou d'un service d'information, joue un rôle central car il constitue une "norme" ou un "standard" pour lesdites activités.

Références bibliographiques

ADAM (N.), BHARGAVA (B.), YESHA (Y.), eds.
1995, *Digital Libraries*, Springer Verlag.

AITCHISON (J.), GILCHRIST (A.)
1992, *Construire un thesaurus*, ADBS Editions.

BARKER (P.)
1991, "An Object Oriented Approach to Hypermedia Authoring", in *Interactive Multimedia Learning Environments*, M. Giardina, ed., Springer.

CHEONG (F. C.)
1996, *Internet Agents : Spiders, Wanderers, Brokers, and Bots*, E. de Pablo (ZfS), New Riders Pub.

[D. E. I. D.]
1997, *Dictionnaire encyclopédique de l'information et de la documentation*, Nathan.

DUPOIRIER (G.)
1995, *Technologie GED*, Hermès.

ECO (U.)
1972, *La Structure Absente*, Mercure de France.

GREIMAS (A.-J.), COURTÉS (J.)
1979, *Sémiotique : dictionnaire raisonné de la théorie du langage*, Hachette-Université.

HJEMSLEV (L.)
1968 (1943), *Prolégomènes à une théorie du langage*, Minuit.

HESLOP (B.), BUDNICK (L.)
1995, *Publier sur le Web : page HTML, Scripts CGI, Formulaire, "Texte-appeal"*, International Thompson Publ.

LAUFER (R.), SCAVETTA (D.)
1992, *Texte, hypertexte, hypermédia*, PUF, 1992.

LUBBICH (H.)
1995, *Towards a CSCW Framework for Scientific Cooperation in Europe*, Springer Verlag.

MUCCHIELLI (A.)
1995, *Les Sciences de l'information et de la communication*, Hachette.

PFEIFFER (H. D.), NAGLE (T. E.)
1993, *Conceptual Structures : Theory and Implementation*, Springer Verlag.

POMART (P. D.)
1997, "Veille", in *Dictionnaire encyclopédique de l'information et de la documentation*, Nathan.

SOWA (J. F.)
1984, *Conceptual Structures*, Addison-Wesley Publishing Compagny.

STOCKINGER (P.)

1992, "Esquisse méthodologique pour la conception de scénarios multimédias", *European Journal for Semiotic Studies*, 4-3, 1992, p. 557-587.

1993, "Multimedia and Knowledge Based Systems", *European Journal for Semiotics Studies*, 5-3, p. 387-424.

TAESCH-WAHLEN (D.)

1997, *Concevoir, réaliser et organiser une bibliothèque*, Ed. du Cercle de la Librairie.

TRAVIS (B.), WALDT (D.)

1995, *The SGML Implementation Guide : A Blueprint for SGML Migration*, Springer Verlag.

RÉSUMÉS

*Savoirs lexicaux
et savoirs encyclopédiques*

t.2 Concevoir et utiliser les hyperdocuments

Patrick J. COPPOCK

L'Encyclopédie comme hypermedia électronique

9-44

Dépasser les contraintes d'un ouvrage de référence

A la lumière des technologies de communication informatisées, cet article propose un examen critique de l'idée convenue qui veut qu'une encyclopédie soit un ouvrage de référence plus ou moins statique et dont l'autorité serait immuable. Les hypermédia électroniques paraissent offrir la possibilité de poser de manière plus constructive le problème de la représentation des recontextualisations incessantes des significations dont la nature et la culture sont porteuses. Avec le développement et l'expansion du *Web*, les processus de changement peuvent être plus facilement (re)présentés comme des processus dynamiques à travers la mise à jour du contenu et de la forme à l'intérieur des documents partagés et des pages-écran. Les problèmes de surcharge informative et de recherche sélective dans les systèmes hypermédia sont envisagés, de même que sont explorés les potentiels et les limites des systèmes de raisonnement de nature *abductive* (Peirce) comme moyen de désambiguïsation. Enfin, le rôle culturel de l'interaction dans la désambiguïsation et le développement du sens est mis au premier plan à travers l'exploration des communautés interprétatives virtuelles, dans le partage conversationnel entre plusieurs utilisateurs qui, en devenant des nœuds dynamiques dans les systèmes hypermédia partagés, lancent un défi à l'autorité d'une encyclopédie envisagée comme un ouvrage de référence canonique.

Davide MUSELLA, Marco PADULA

Extraire l'information pertinente des larges répertoires électroniques

129-150

Le monde de l'information tel qu'il est représenté par Internet n'est pas totalement défini en raison de son évolution rapide. Cet article apporte une vue controversée des problèmes qui s'y posent, des orientations destinées à leur résolution, et propose en conséquence quelques définitions. Il fait le choix de tracer le cadre de l'activité du spécialiste de l'information et de son environnement ; celui-ci doit abandonner dans son approche idées, scénarios, considérations philosophiques et définitions préconçus et doit maintenir un lien entre une attitude purement spéculative et le pragmatisme. Alors qu'il analyse les besoins sociaux, il conçoit

des solutions en fonction des infrastructures disponibles, ou prévisibles, et tient compte des méthodologies existantes ainsi que des langages de la communication.

David PIOTROWSKI

Lexicographie et formes opératoires de l'hypertextualité 65-102

Cette étude présente une analyse comparative de trois modèles de l'hypertextualité : (i) le modèle de Dexter, (ii) le modèle "à 3 modules" et (iii) le modèle "à 2 Couches". Le modèle de Dexter, qui définit un cadre générique des formes hypertextuelles, envisage les hypertextes sous l'angle exclusif de leur structure fonctionnelle interne, statique ou dynamique. A la différence du modèle de Dexter, le modèle "à 3 modules" alloue une place explicite au lecteur : il permet de prendre en considération l'activité fonctionnelle de lecture. Dans le modèle à "2 couches", la structure fondamentale de l'hypertextualité est une structure d'appropriation récursive : un réseau de qualifications projeté sur les textes oriente les lectures qui conduisent à l'élaboration de nouveaux systèmes de qualifications. On montre que les trois modèles se distribuent suivant un double gradient : le premier paramétré par la précision d'analyse des formes internes de l'hypertextualité, le second par la largeur de la perspective englobant l'activité de lecture. Ainsi, ces trois modèles ne sont pas mutuellement exclusifs : ils offrent des images plus ou moins globales, plus ou moins focalisées d'un même fait empirique : l'hypertextualité. C'est donc dans leur combinaison que réside une problématique satisfaisante de l'hypertextualité.

**Elisabeth STOCKINGER de PABLO, Hammou FADILI,
Peter STOCKINGER**

SemioNet

Spécification, production et implémentation de services d'information en ligne 151-194

Les développements des technologies de la communication et de l'information concernant la mise en réseau intéressent de plus en plus les secteurs de la recherche et de l'éducation. Les applications les plus connues sont l'enseignement à distance, l'édition en ligne ou encore le travail collectif. Or la concrétisation de ces applications nécessite aussi bien une infrastructure informatique adéquate qu'une réelle réflexion sur la forme et le fond des documents et services à produire. Aussi les problèmes de modélisation, de conception, de spécification, de production et d'implémentation de tels documents et services, ainsi que leurs problèmes de maintenance, ne sont pas tellement d'ordre technologique mais plutôt d'ordre sémiotique. A partir d'un exemple concret, les auteurs de cet article essaient de démontrer que l'approche sémiotique du document peut constituer un cadre d'intégration des différentes méthodologies locales existantes et former ainsi la base d'une méthodologie systématique et générale appelée *SémioNet*. *SemioNet* est conçu pour guider les différentes phases de construction, de réalisation et de réutilisation de documents et services à support électronique.

Chris TOMPSETT*L'Education, les systèmes d'information
et les structures de connaissances*

103-128

Etudier au moyen des systèmes d'information est à mettre en relation avec le potentiel croissant des ressources électroniques pour la conception et la fourniture de matériels d'éducation et de formation. L'accent est mis sur la délivrance d'outils qui permettent à l'étudiant d'exploiter efficacement cette richesse, car il est nécessaire que les techniques d'accès développées soient plus directes que les approches globales de recherche d'information et moins contraintes que celles validées individuellement. L'analyse des bases de connaissances offre des possibilités : un prototype a été développé qui génère automatiquement les liens entre les nœuds hypertexte utilisant une description des connaissances à partir du formalisme KL-One. L'évaluation d'un tel système et l'impact des récents changements sur les bases de connaissances y sont décrits.

**William A. TURNER, Boris BORZIC,
Mathilde de SAINT-LEGER, Yves MULLER***La Sémiotique documentaire*

45-64

La sémiotique documentaire appliquée à la littérature scientifique est susceptible de contribuer à l'organisation et à l'optimisation des collaborations scientifiques à condition de s'affranchir des perspectives logiciste (le langage envisagé comme une copie de la réalité) et sociologiste (le langage considéré comme l'expression d'un comportement social dicté par des valeurs, des motivations). Ce texte adopte une approche économiste afin d'étudier la façon dont les interactions sociales s'organisent autour de l' "échange" de ces "objets" que sont les publications scientifiques. L'analyse sémiotique des flux documentaires permet alors de faire émerger les centres d'intérêt thématiques propres à une communauté scientifique donnée, et surtout les formes de collaboration (partenariat, sous-traitance, alliance) envisageables entre ses membres. Le cas pratique traité ici est celui de publications abordant le thème des transformations sociales dans la ville.

