
Du virtuel au manipulable

Quelles interfaces homme/machine pour l'appropriation de documents dans les bibliothèques numériques ?

Mona Farhat — Aurélien Bénel

*ISTIT/Tech-CICO
Université de Technologie de Troyes
12 rue Marie Curie
10000 Troyes Cedex
techcico@utt.fr*

RÉSUMÉ. Les bibliothèques numériques ont souvent été vantées pour leur "virtualité". L'accès au livre "dématérialisé" est en effet quasi-immédiat, mondial et en autant d'exemplaires que nécessaire. Reste à inventer une certaine matérialité du document numérique qui permettrait, au même titre que le document "papier", d'être le support des gestes du lecteur ; ces gestes d'appropriation : photocopier, annoter, classer... Dans ce but, notre équipe a mis au point des modèles de données dans le cadre de travaux antérieurs. Dans cet article, nous nous focalisons sur le modèle d'activité de l'appropriation de documents afin, d'une part, d'expliquer un certain nombre de problèmes ergonomiques rencontrés et, d'autre part, de voir comment des systèmes hypermédias marquants ont pu les résoudre. Enfin, nous faisons une proposition d'interface homme/machine pour ce genre d'outils d'appropriation de documents numériques.

MOTS-CLÉS : Bibliothèques numériques, Annotation, Appropriation, Ergonomie des systèmes hypermédias, Modèles d'activité, Hypermédias collaboratifs.

1. Introduction

Les bibliothèques numériques ont souvent été vantées pour leur "virtualité". L'accès au livre "dématérialisé" est en effet quasi-immédiat, mondial et en autant d'exemplaires que nécessaire. La lecture dans ses bibliothèques numériques semble pourtant plus difficile, comme si la compréhension des ouvrages nécessitait une « com-préhension » (prise avec soi) du volume matériel. Reste donc à inventer une certaine matérialité du document numérique qui selon nous permettrait, au même titre que le document "papier", d'être le support des gestes d'appropriation du lecteur.

Dans une première partie, nous tâcherons de mieux définir ce que nous appelons « appropriation », et de voir comment se traduit cette activité dans les bibliothèques traditionnelles.

Le modèle de cette activité d'appropriation nous permettra, dans une deuxième partie, d'analyser un certain nombre de systèmes hypermédias, y compris un système développé dans nos travaux antérieurs : *porphyry*.

Dans une troisième partie, nous tenterons d'expliquer un certain nombre de problèmes ergonomiques rencontrés, et de tenir compte des systèmes hypermédias étudiés précédemment, pour construire une proposition d'interface homme/machine pour ce genre d'outils d'appropriation de documents numériques.

2. Appropriation de documents traditionnels

Afin de réinventer cette matérialité perdue du document numérique, un certain nombre de chercheurs se sont tournés vers les usages (observés ou prescrits) du document traditionnel.

Ainsi, Kenton O'Hara et son équipe (O'Hara *et al.*, 1998), dans le but de concevoir les interfaces homme/machine des futures bibliothèques numériques, ont observé l'activité des doctorants dans une bibliothèque traditionnelle. Il en ressort une grande diversité de tâches :

- la lecture,
- l'annotation,
- la rédaction de fiches de lecture,
- la reproduction par photocopie,
- la consultation d'une bibliographie,
- la relecture,
- l'écriture de nouveaux documents (articles, mémoires, bibliographies...).

La transposition dans le monde numérique de l'ensemble de ces tâches et la convergence des interfaces associées seront sans doute nécessaires si l'on souhaite

obtenir un jour une réelle bibliothèque numérique (Cubaud, 2005). Cependant, nous nous focaliserons ici sur ce qui caractérise au moins quatre de ces tâches : la *capitalisation de traces matérielles*. Nous l'appellerons « appropriation »¹.

Cette thématique se retrouve d'ailleurs au coeur des travaux de Jacques Virbel et de son équipe (Mazhoud *et al.*, 1995) autour de la conception pour la Bibliothèque nationale de France [BNF] de Postes de lecture assistée par ordinateurs [PLAO]. Ces travaux en collaboration avec les lecteurs permirent de mettre en évidence huit verbes d'action :

- Hiérarchiser (attribuer un niveau d'importance),
- Architecturer (typer des fragments),
- Contextualiser (borner le sens),
- Reformuler,
- Commenter (critique, idée associée),
- Documenter (ajouter un fragment pour en comprendre une autre),
- Corréler (relier deux fragments),
- Programmer (projeter une action : à traduire, à relire, à analyser...).

Ces actions, pourtant inscrites dans notre tradition interprétative forment une activité globale difficile à qualifier avec des mots courants. L'équipe parle de *lecture active*, ou, dans un sens élargi, de *glose* ou d'*annotation* (Veron, 1997).

Plus récemment, dans le cadre des travaux de thèse de Gaëlle Lortal, un modèle d'activité a été proposé pour la conception d'un logiciel d'annotation collective (Lortal *et al.*, 2005). Ce modèle repose sur les méthodes médiévales de production, à partir d'une source documentaire, d'un commentaire critique. Il distingue les actions suivantes :

- *Cogitatio* (sélection dans une source documentaire de fragments),
- *Collatio* (regroupement de fragments documentaires dans des collections),
- *Inventio* (organisation de différentes collections sous forme d'un plan),
- *Dictamen* (rédaction d'un brouillon à partir d'un plan),
- *Exemplar* (soumission d'une copie à un supérieur ou à des pairs),
- *Emendare* (publication de la version finale).

Ce troisième modèle de l'activité d'appropriation de documents traditionnels nous semble particulièrement opératoire pour la conception de logiciels. Comme nous allons le voir, la plupart des actions recensées ont d'ailleurs déjà fait l'objet de tentatives d'implémentation dans le domaine de l'Hypermédia.

¹ Le mot « appropriation » a donc davantage ici le sens proposé par Andrea Iacovella (Iacovella, 2002) que celui proposé par Pierre Rabardel.

3. Etat de l'art dans les systèmes hypermédias

A l'aide du modèle d'activité choisi, nous allons dans cette partie analyser comment les différentes actions ont été implémentées à travers un petit panel de systèmes hypermédias :

– *Hyper-G* (Maurer, 1996) et son client graphique *Harmony* (Andrews, 1996) de l'université technologique de Graz (Autriche), devenus aujourd'hui un produit commercial de « gestion coopérative de contenus » appelé « *HyperWave* »,

– *ATLAS.ti*® (Muhr, 1997), issu d'un projet de recherche du département de psychologie de l'Université Technique de Berlin, aujourd'hui commercialisé en tant qu'« atelier d'analyse qualitative de documents »,

– *TheBrain*® (TheBrain, 2001), inventé par un jeune autodidacte en informatique, aujourd'hui protégé par plusieurs brevets et commercialisé en tant que « plateforme de connaissance »,

– *CritLink* (Yee, 2002), serveur d'« annotation publique du Web » développé par le Foresight Institute, librement utilisable il y a quelques années,

– *Porphyry* (Bénel, 2004), développé dans le cadre du réseau interdisciplinaire ARTCADHi², aujourd'hui diffusé en open-source en tant qu'« outil de lecture-écriture pour experts »,

– *LUCIA* (Perlerin, 2004), ensemble d'outils de « recherche d'information personnalisée³ » développé au sein du laboratoire GREYC à Caen, et dont le modèle est basé sur la sémantique interprétative de François Rastier⁴.

Concernant la sélection de fragments, on peut tout d'abord constater qu'elle est toujours dépendante du type de source : positions de début et de fin pour les textes, forme 2D plus ou moins complexe pour les images (*porphyry*, *ATLAS.ti*, *Hyper-G*), forme 2D animée pour la vidéo (*Hyper-G*), forme 3D pour les scènes de réalité virtuelle (*Hyper-G*). On notera également qu'un système comme *LUCIA* permet de distinguer différents ensembles de fragments à l'aide de couleurs. Enfin, on retiendra que certains systèmes tentent de rendre visible le chevauchement de fragments (*LUCIA*, *CritLink*, *ATLAS.ti*®), en particulier à l'aide de barres marginales (*ATLAS.ti*®).

² ARTCADHi, Atelier de recherche transdisciplinaire pour la construction du sens dans les disciplines historiques, animé par Andrea Iacovella, <<http://www.porphyry.org/>>.

³ Il s'agit donc davantage d'un outil de visualisation que de manipulation. Cependant les métaphores utilisées pourraient tout à fait être utilisées pour la manipulation.

⁴ *LUCIA* est comparable en cela à *PASTEL* (Tanguy, 1998) : « Programme d'Aide à l'Analyse Sémantique de Textes, même Littéraires », développé au sein de l'ENST-Br de Brest.

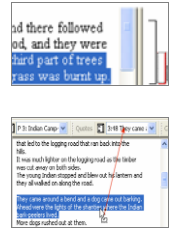
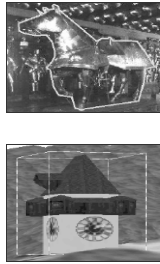

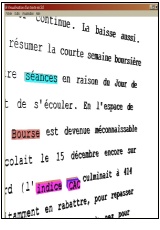
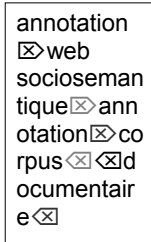
ATLAS.ti®	Hyper-G	Porphyry	LUCIA	CritLink
				

Tableau 1. Sélection dans une source documentaire de fragments (état de l'art)

Concernant le regroupement de fragments documentaires dans une collection, nous retiendrons l'idée de disposer de deux affichages (*ATLAS.ti*®), l'un en indiquant les collections en marge de la source, l'autre en extrayant pour une collection donnée tous les fragments.

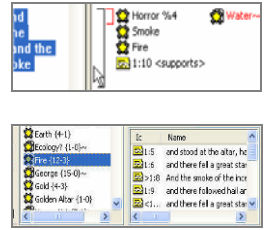
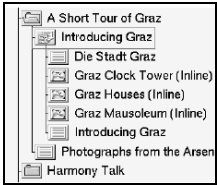
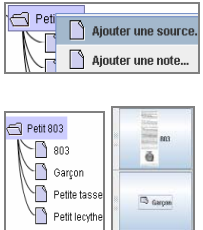
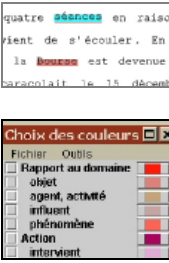
ATLAS.ti®	Hyper-G	Porphyry	LUCIA
			

Tableau 2. Regroupement de fragments documentaires dans des collections (état de l'art)

Concernant l'organisation des collections, le mode de visualisation est en général assez dépendant du modèle sous-jacent. Ces modèles sont d'ailleurs assez divers : graphe orienté acyclique (*Porphyry*, *Hyper-G*, *ATLAS.ti*®), graphe orienté non contraint avec liens typés (*TheBrain*®), modèle bi-dimensionnel (*LUCIA*). *Porphyry* propose par ailleurs la navigation conjointe entre plusieurs « points de vue ».

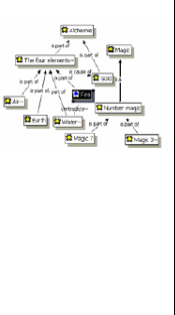
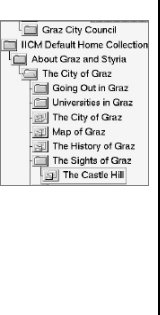
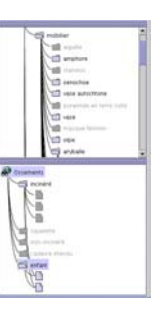

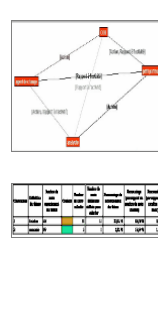
ATLAS.ti®	Hyper-G	Porphyry	The Brain®	LUCIA
				

Tableau 3. Organisation de différentes collections sous forme d'un plan (état de l'art)

Pour la phase de rédaction du brouillon à partir d'un plan, tous les systèmes étudiés lancent pour chaque « îlot » de texte une fenêtre d'édition.

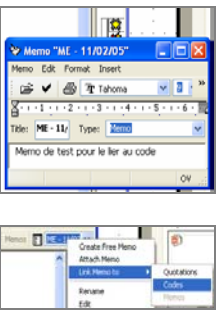
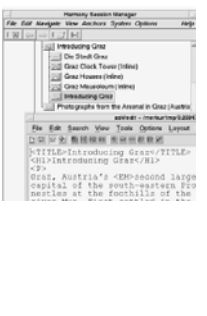


ATLAS.ti®	Hyper-G	Porphyry	The Brain®
			

Tableau 4. Rédaction d'un brouillon à partir d'un plan (état de l'art)

Les actions de soumission et de publication sont en général assez mal gérées. Les systèmes oscillent entre le totalement privé (monoposte) et le totalement public. Certains mettent en place des politiques minimales de droits d'accès. Nous travaillons depuis peu sur une « gestion des identités » qui permettrait de mieux prescrire et décrire l'activité des sujets connaissant dans ce type de systèmes. Nous ne pourrions pas aborder cette thématique dans les limites cet article.

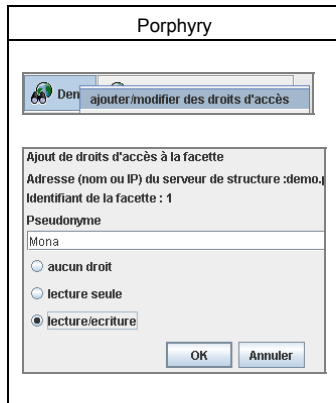


Tableau 5. *Soumission d'une copie à un supérieur ou à des pairs, publication de la version finale (état de l'art)*

4. Proposition d'interface pour l'appropriation de documents numériques

Comme le montrent particulièrement bien les interfaces d'*ATLAS.ti*® et de *LUCIA*, l'*organisation des collections* doit se situer dans un tout autre espace de travail que celui de la *sélection de fragments* et du *regroupement de ces fragments en collections*. En effet, l'objet de référence diffère : dans le premier cas, il s'agit d'un corpus documentaire, quand, dans le second, il s'agit d'une source documentaire unique. Il est ainsi souhaitable de disposer de deux fenêtres, l'une centrée sur le corpus, l'autre centrée sur la source, plutôt que de disposer d'une « interface multi-document » [MDI] comme c'est le cas dans *porphyry*. Il est même préférable de disposer de deux applications différentes, de sorte que l'outil de gestion de corpus permette de gérer n'importe quel type de documents sources. Lors de l'ouverture d'une source, il est alors possible de lancer, pour les formats reconnus, l'outil de gestion de source et, pour les autres, des applications tierces.

Pour devenir cet outil de gestion de corpus, *porphyry* doit seulement subir quelques modifications. Tout comme, sur une étagère, seuls apparaissent les livres et non les annotations qui sont à l'intérieur, de même, les fragments ne doivent pas apparaître dans le corpus. En retour, les collections qui les décrivent doivent apparaître comme décrivant les sources dont ils sont issus.

Reste à inventer ce que pourrait être cet outil de gestion de source documentaire. De cet outil, que nous appellerons provisoirement « *laSuli* » (logiciel d'annotation sociale à l'usage des lecteurs-interprètes), nous donnerons une maquette sous forme de trois exemples d'utilisation (fig.1-3).

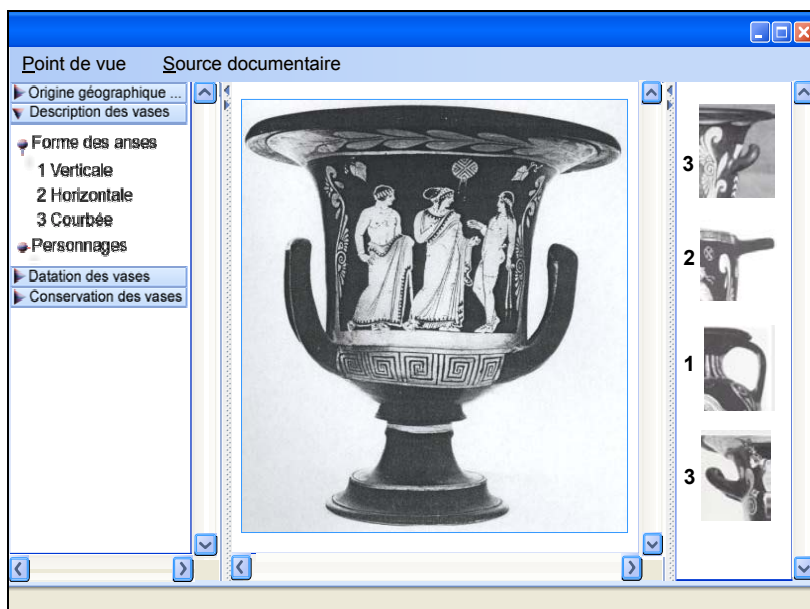


Figure 1. *LaSuli (maquette) : Analyse typologique d'un vase d'après photographie*

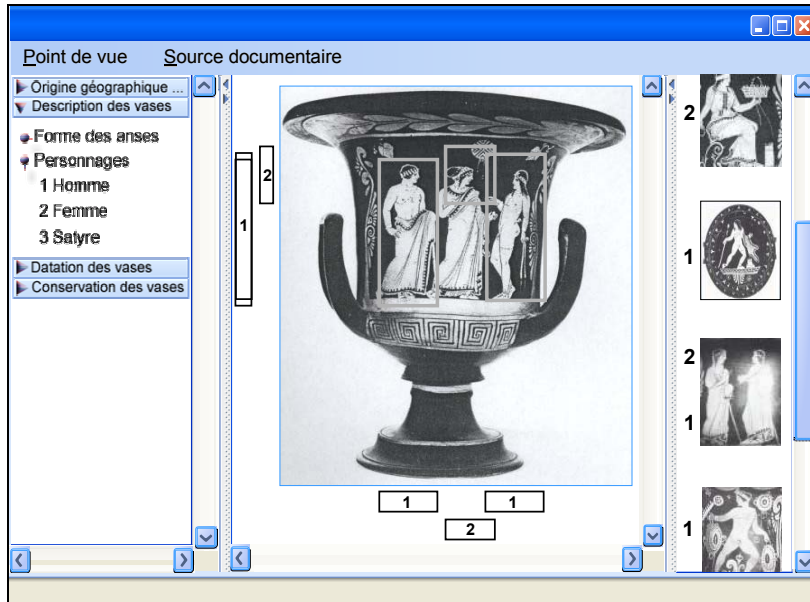


Figure 2. *LaSuli (maquette) : Analyse iconographique d'un vase d'après photographie*

Dans un système comme *LUCIA*, basé sur une sémantique différentielle, seuls deux niveaux de collections sont disponibles : ce qui rassemble et ce qui différencie. Ce type de structure, peu adapté pour un corpus, est par contre redoutablement efficace pour l'étude d'une source. Ainsi, contrairement à *porphyry*, il serait souhaitable que *laSuli* n'affiche à chaque instant que deux niveaux de collections (Fig. 1-2). Cela n'empêcherait d'ailleurs nullement l'utilisateur de changer de niveau quand il le souhaite.

Porphyry permet d'extraire un fragment. Cependant, cette fonctionnalité, proposée comme une sorte de « zoom » de la source, ne rencontre que peu d'intérêt auprès des usagers. Il est intéressant de noter qu'*ATLAS.ti*® utilise l'extraction de fragments (de texte) pour rassembler à l'écran les fragments d'une même collection. En développant cette idée à l'aide de la sémantique différentielle (à la base de *LUCIA*), on pourrait imaginer d'utiliser dans *laSuli* les fragments de différentes collections pour les comparer avec la source (Fig. 1) ou avec des fragments de la source (Fig. 2)

Porphyry matérialise sur une source les fragments d'une seule collection. *ATLAS.ti*® quant à lui montre l'ensemble des fragments définis sur une source. Pour

laSuli, nous préconiserions là encore une approche différentielle du problème, à savoir montrer les fragments d'une collection et les différencier en fonction des sous collections (Fig.2-3).

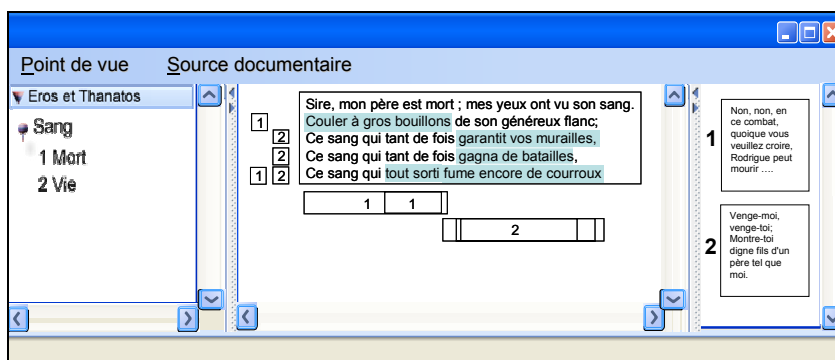


Figure 3. *laSuli* (maquette) : Commentaire composé d'un texte du Cid

Un dernier aspect concerne la nécessité de manipuler les objets directement dans leur contexte. Ainsi, plutôt que d'ouvrir une fenêtre pour créer un fragment, il serait préférable dans *laSuli* de le faire directement à partir de la source (comme dans *ATLAS.ti*® et *Hyper-G*). De même, il serait souhaitable de modifier directement dans leur contexte les notes dynamiques (comme dans *TheBrain*®). En effet, c'est en contexte que les notes et les fragments prennent sens.

5. Bilan et perspectives

Au cours de cet article, nous avons exploré la question de l'appropriation de documents numériques. A l'issue de la première partie, nous avons adopté un modèle d'activité issu des méthodes médiévales de production d'un commentaire critique à partir d'une source documentaire. Dans la deuxième partie, nous avons étudié un certain nombre d'outils, en particulier : *ATLAS.ti*®, assez abouti du point de vue ergonomique pour la gestion des sources documentaires ; *LUCIA*, offrant des visualisations basées sur la sémantique interprétative de François Rastier, et enfin, *porphyry*, notre outil, spécialisé dans la gestion des corpus. Dans une troisième partie, nous avons tenté de répondre aux problèmes ergonomiques rencontrés avec *porphyry* dans la gestion des sources documentaires. Pour cela, nous avons été amenés à concevoir l'interface homme-machine d'une application, complémentaire à *porphyry*, appelée « *laSuli* ».

Une fois développée, cette application sera mise à l'épreuve de deux manières : tout d'abord, sous la forme d'expérimentations ponctuelles avec des groupes

d'étudiants, ensuite, par sa mise en service, en particulier auprès des experts qui utilisent déjà *porphyry* dans leur activité.

Remerciements

Les recherches présentées ont été co-dirigées par Myriam Lewkowicz, les auteurs tiennent à la remercier chaleureusement. Les auteurs remercient également Nadia Gauducheau et François Helly pour leurs remarques expertes sur les premières maquettes, ainsi que Rami El Sawda pour sa précieuse participation. Enfin, les auteurs remercient l'équipe du CRATA de Toulouse pour leurs retours sur l'usage de *porphyry* dans leurs activités de recherche sur la typologie et l'iconographie des vases.

6. Bibliographie

- Andrews K., *Browsing, building, and beholding cyberspace: New approaches to the navigation, construction, and visualisation of hypermedia on the Internet*, Technical Sciences Doctor dissertation, Graz University of Technology, 1996. Disponible sur : <<http://www2.iicm.edu/keith-phd/>>.
- Bénel A., Consultation assistée par ordinateur de la documentation en Sciences Humaines : Considérations épistémologiques, solutions opératoires et applications à l'archéologie, In : F. Rastier (Ed.), *Texte ! mars 2004*. Disponible sur : <<http://www.revue-texto.net/lnedits/Benel/Benel.html>>.
- Cubaud P., Interaction 3D pour les bibliothèques numériques, In : F. Papy (Ed.), *Les bibliothèques numériques*, Paris : Hermès, 2005.
- Iacovella A., Modèle opératoire de navigation pour les experts : Appropriation sémantique et délimitation de l'espace documentaire, In : F. Ghitalla (Ed.), *La navigation (Numéro spécial), Les cahiers du Numérique, vol.3, n°3*, Paris : Hermès, 2002. p. 175-190. Disponible sur : <http://www.porphyry.org/Members/aiacovella/iacovella_LCN_02.pdf>.
- Lortal G., Lewkowicz M., Todiracu-Courtier A., Modélisation de l'activité d'annotation discursive pour la conception d'un collectif support à l'herméneutique, In : *Actes des 16^{ème} journées francophones d'ingénierie des connaissances [IC 2005]*, Grenoble : PUG, 2005. p. 169-180.
- Maurer H. (Ed.), *HyperWave: The Next Generation Web Solution*, Addison-Wesley, 1996, 635 p. Disponible sur : <<http://www.iicm.edu/hwbook/>>.
- Mazhoud O., Pascual E., Virbel J., Représentation et gestion d'annotations, In : *Hypertextes et hypermédias : Réalisations, outils et méthodes*, Techniques de l'Information, Paris : Hermès, 1995. p. 127-138.
- Muhr T., *Atlas/ti: Short User's Manual*, Berlin : Scientific Software Development, 1997.
- O'Hara K., Smith F., Newman W., Sellen A., Student readers' use of library documents: implications for library technologies, In: *Conference Proceedings on Human Factors in*

Computing Systems [CHI'98], New York : ACM Press, 1998. p. 233-240. Disponible sur : <<http://doi.acm.org/10.1145/274644.274678>>.

Perlerin V., Sémantique légère pour le document : Assistance personnalisée pour l'accès au document et l'exploration de son contenu, In : F. Rastier (Ed.), *Texto ! décembre 2004*. Disponible sur : <<http://www.revue-texto.net/Inedits/Perlerin/Perlerin.html>>.

Personal Brain: User Guide, TheBrain Technologies Corporation, 2001, 99 p. Disponible sur : <http://www.thebrain.com/products/personalbrain/support/Manual/PB_User_Guide.pdf>

Tanguy L., Traitement automatique de la langue naturelle et Interprétation : Contribution à l'élaboration d'un modèle informatique de la Sémantique Interprétative, In : F. Rastier (Ed.), *Texto ! mars 1998*. Disponible sur : <http://www.revue-texto.net/Inedits/Tanguy/Tanguy_these.html>.

Veron M., *Modélisation de la composante annotative dans les documents électroniques*, Mémoire de DEA RCFR INPT, Toulouse, Septembre 1998. Disponible sur : <http://www.enseeiht.fr/fr/recherche/info/Intellig/PUBLICATIONS/VERON/dea_mv.txt>.

Yee K.-P., CritLink: Advanced hyperlinks enable public annotation on the web, Demonstration abstract, In : *ACM Conference on Computer-Supported Cooperative Work, New Orleans, November 16-20, 2002 [CSCW'02]*. Disponible sur : <<http://zesty.ca/crit/yee-crit-cscw2002-demo.pdf>>.