



HAL
open science

Vers un outil documentaire unifié pour les chercheurs en archéologie

Aurélien Bénel, Sylvie Calabretto, Jean-Marie Pinon, Andrea Iacovella

► To cite this version:

Aurélien Bénel, Sylvie Calabretto, Jean-Marie Pinon, Andrea Iacovella. Vers un outil documentaire unifié pour les chercheurs en archéologie. 18e congrès INFORSID, May 2000, Lyon, France. hal-02938667

HAL Id: hal-02938667

<https://utt.hal.science/hal-02938667>

Submitted on 18 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vers un outil documentaire unifié pour les chercheurs en archéologie

Aurélien BENEL^{*,+}, Sylvie CALABRETTO^{*},

Andréa IACOVELLA⁺, Jean-Marie PINON^{*}

* Laboratoire d'Ingénierie des Systèmes d'Information, INSA de Lyon,
Bâtiment 501, 20 avenue Albert Einstein, 69621 Villeurbanne Cedex, France.
{Aurelien.Benel, Sylvie.Calabretto, Jean-Marie.Pinon}@lisi.insa-lyon.fr

Tél. : (33) 4 72 43 84 81, Fax : (33) 4 72 43 85 18.

+ Ecole française d'Athènes, 6 rue Didotou, 106 80 Athènes, Grèce.
Andrea.Iacovella@efa.gr Tél. : (30) 1 3612 521, Fax : (30) 1 3631 849.

[Catégorie Chercheur]

Résumé :

Nous nous situons dans le contexte d'un corpus de documents adressé à des chercheurs en archéologie et proposons de bâtir des outils facilitant leurs tâches documentaires. Nous introduisons une approche mettant l'accent sur l'activité humaine, l'aspect exploratoire de cette activité, ainsi que son caractère contextuel (corpus et actions précédentes). Nous proposons un modèle d'indexation (graphe acyclique) et un algorithme de filtrage permettant d'assister l'être humain autant dans ses activités de recherche d'information que dans celles d'indexation, de lecture d'un document structuré, d'annotation et de partage de notes.

Mots-clés :

Recherche d'information, Indexation, Réseaux sémantiques, Lecture/Annotation, Collaboration, Navigation.

Abstract:

This paper deals with the case of researchers in archaeology working with a corpus of documents. We aim at building methods and tools in order to make their documentary tasks easier. Our approach emphasizes the human, exploratory and contextual aspects of this activity. The indexing model (acyclic graph) and filtering algorithm we propose provides assistance in information retrieving, indexing, structured documents reading, annotating and note sharing.

Keywords:

Information retrieval, Indexing, Semantic nets, Reading/annotating, Collaboration, Navigation.

1 Contexte de l'étude : la Chronique des fouilles de l'EFA

L'Ecole française d'Athènes, établissement public de recherche dans les disciplines se rapportant au monde grec assure la diffusion de ses recherches grâce à ses publications. L'une d'elle, la *Chronique des fouilles* (au sein du *Bulletin de Correspondance Hellénique*), a pour mission de signaler aux lecteurs toutes les « nouveautés » archéologiques en Grèce et à Chypre. Le corpus, formé de l'ensemble de ses numéros annuels de 1920 à nos jours, constitue une source d'information incomparable pour les chercheurs en archéologie.

C'est dans ce contexte d'un corpus spécialisé adressé à des chercheurs que s'inscrit notre travail visant à offrir des outils informatiques pouvant faciliter leurs tâches documentaires. Nous commencerons donc par identifier le type de tâches documentaires effectuées par ces chercheurs, pour ensuite adopter une approche de la recherche d'information qui y soit cohérente, et enfin bâtir sur ces bases un modèle permettant à l'ordinateur d'assister au mieux le chercheur dans ces tâches.

2 Activités documentaires

Notons bien qu'il serait vain, dans un domaine nécessitant une expertise de si haut niveau et une si totale liberté, de prétendre établir un « modèle du domaine » à l'image de ce que l'on peut obtenir en informatique de gestion. Cependant, on peut tenter de décrire quelques activités documentaires typiques des chercheurs.

Kenton O'Hara et son équipe (cf. [OHAR98]) écrivaient que les tâches exercées dans les bibliothèques universitaires allaient au delà de la traditionnelle recherche d'information. En effet, la lecture, l'annotation, la rédaction de fiches de lecture, l'action de photocopier, la consultation et l'enrichissement d'une bibliographie, la relecture et enfin l'écriture de nouveaux documents (articles, mémoires...) devraient être considérés comme formant un tout.

Klaus Tochtermann (cf. [TOCH94]), quant à lui, prenait position pour que les interactions sociales ne soient pas oubliées dans les bibliothèques virtuelles. En effet, comme il le rappelait, les bibliothèques ne sont pas uniquement des collections de documents mais aussi des lieux de communication, voire de collaboration. Même en se limitant à l'aspect de la recherche de documents, l'interrogation (directe ou par l'intermédiaire de bibliographies, dossiers, index...) de spécialistes (bibliothécaires ou autres chercheurs) reste indispensable. Eux seuls sont capables de comprendre le contenu des ouvrages, de porter dessus un jugement critique, d'en faire une synthèse, de faire des recoupements avec d'autres ouvrages, de les compléter par des informations non-officielles ou subjectives, de comprendre la question du chercheur d'information et de l'aider si besoin à « raffiner » ou à étendre sa requête.

Dans l'esprit des deux articles évoqués ci-dessus, on peut répertorier un certain nombre d'activités exercées par les chercheurs en présence d'un corpus comme la *Chronique des fouilles* :

- la recherche d'un fragment de document en fonction d'un certain nombre de critères croisés (lieu de la fouille, datation, type d'objets trouvés, directeur...),
- la recherche d'un fragment de document relié explicitement à un fragment donné (par référence bibliographique ou par partage d'une partie de leur hiérarchie des titres),
- la recherche d'un fragment de document relié implicitement à un fragment donné (relation mise en évidence par leur présence commune dans une même bibliographie...),

- l'indexation officielle « pour le compte » d'une bibliothèque ou du périodique,
- la rédaction d'une fiche de lecture (résumé, extraits, étude critique...),
- la constitution d'une bibliographie sur un sujet donné,
- la communication d'une bibliographie et de son appareil critique (de manière publique dans un article ou privée entre un professeur et un étudiant),
- l'échange de références bibliographiques entre collègues.

3 Quelle approche adopter pour la recherche de documents ?

Au vu de la variété des approches et des modèles existant pour la recherche de documents (cf. [CHIA99]), on peut se demander laquelle adopter. A cette question, Gerard Salton (cf. [SALT68]) répondait par une approche ensembliste. En effet, la recherche de documents traite de deux ensembles : celui des documents (appelé corpus documentaire) et celui des descripteurs (liste « d'autorités » ou liste de mots-clefs construite au fur et à mesure). On peut dès lors construire le graphe (en treillis) de l'ensemble, muni de la relation d'inclusion, de tous les sous-corpus possibles (cf. Fig.1), ainsi que celui (qui sans les « boucles réflexives » serait également un treillis) de l'ensemble des requêtes, muni de la relation de conjonction avec un descripteur (cf. Fig.2).

Or, à partir de la matrice faisant correspondre un document à des descripteurs (cf. Tab.1), il est possible d'en déduire la correspondance existant entre chacune des requêtes possibles et le corpus donné en réponse (cf. Tab.2). On remarque dès lors que, d'une part, un certain nombre de

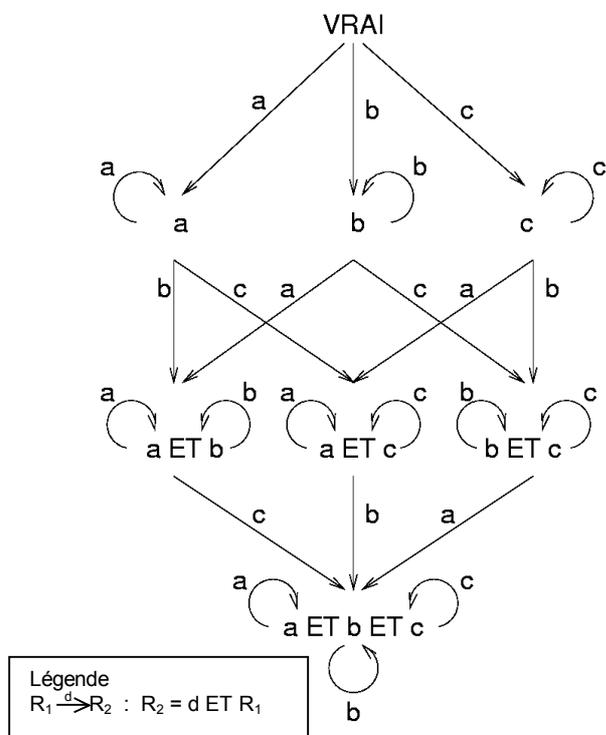
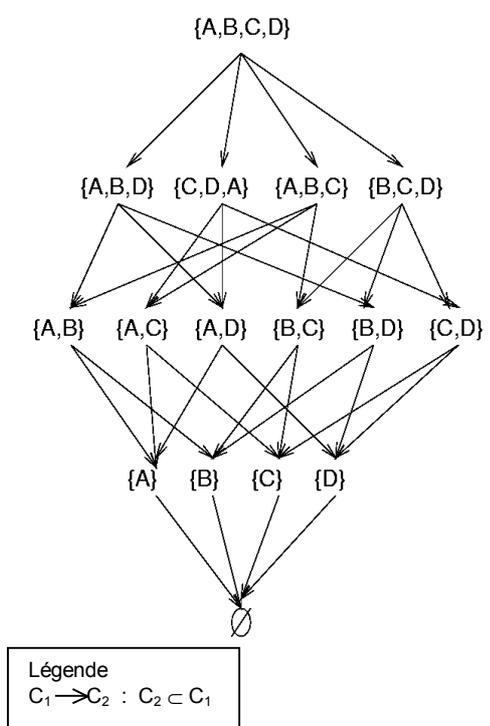


Figure 1 – Graphe d'inclusion des corpus documentaires pouvant être obtenus à partir des documents A, B, C et D.

Figure 2 – Graphe des requêtes pouvant être obtenues par conjonction à partir des descripteurs documentaires a, b et c.

| | | Descripteurs | | |
|-----------|---|--------------|---|---|
| | | a | b | c |
| Documents | A | X | | X |
| | B | X | X | |
| | C | | | X |
| | D | X | | |

Tableau 1 – Index des documents A, B, C, D à l'aide des descripteurs a, b, c.

| VRAI | {A,B,C,D} |
|-------------|-----------|
| a | {A,B,D} |
| b | {B} |
| c | {A,C} |
| a ET b | {B} |
| b ET c | ∅ |
| c ET a | {A} |
| a ET b ET c | ∅ |

Tableau 2 – Correspondance entre requêtes et corpus (cf. Tab.1).

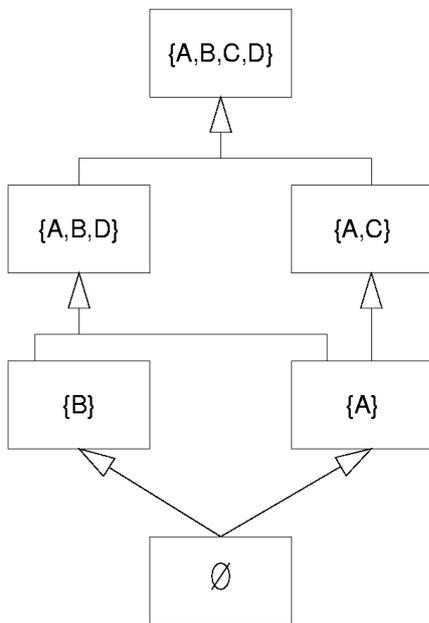


Figure 3 – Diagramme des classes de documents (notation UML cf. [LAI97]) pouvant être obtenues à l'aide des requêtes du Tableau 2 .

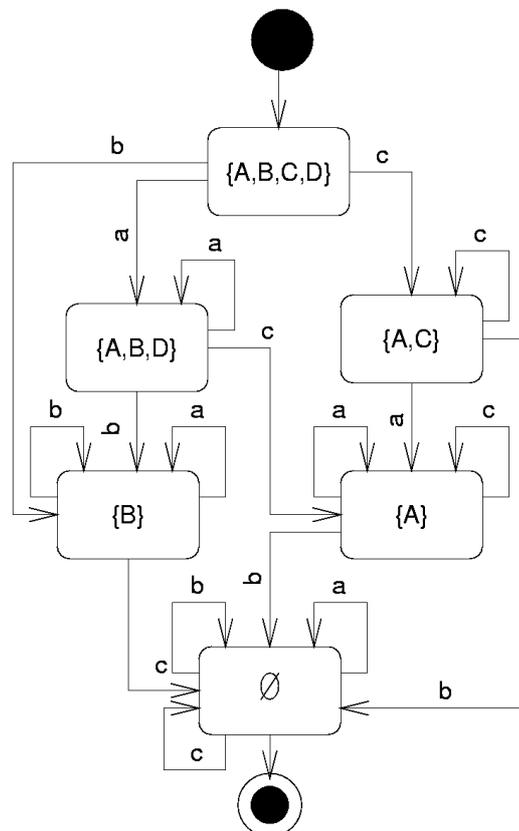


Figure 4 – Diagramme des états de la navigation entre corpus documentaires (notation UML cf. [LAI97], par sélection successive des descripteurs a, b et c (cf. Tab.2).

corpus possibles ne peuvent être obtenus par aucune requête (ex : $\{B,C\}$) et que, d'autre part, plusieurs requêtes possibles peuvent donner en réponse le même corpus (ex : la requête *a ET b* et la requête *b*). De là, on peut :

- soit supprimer du graphe des corpus les corpus inaccessibles : on obtient alors un diagramme de généralisation/spécialisation des classes de documents (cf. Fig.3) pouvant être définies en intention par leurs descripteurs communs (cf. [CARP94]),
- soit fusionner, dans le graphe des requêtes, les requêtes décrivant le même corpus documentaire : on obtient alors un diagramme d'états (cf. Fig.4) dans lequel les états sont des corpus et les transitions des requêtes élémentaires. Ces requêtes à un seul descripteur portant sur des corpus intermédiaires peuvent aussi être vues comme l'ajout de ce descripteur à la requête globale (portant sur le corpus initial) : on parlera d'ailleurs par la suite de « raffinement de la requête » (de l'anglais « query refining »).

C'est cette dernière approche qui retiendra notre attention. En effet, elle met l'accent sur l'activité humaine et rend bien d'une part l'aspect exploratoire de l'activité, d'autre part son caractère contextuel (c'est à dire ici dépendant du corpus précédemment sélectionné).

4 Proposition de modèle

Tâchons maintenant d'étudier comment l'approche que nous venons de définir permet d'établir un modèle unifié pour la recherche, l'indexation et l'annotation (coopérative) de documents (structurés ou non) .

4.1 Recherche de documents

Dans le diagramme introduit précédemment (cf. Fig.4), on remarque que pour un état donné :

1. certaines transitions mènent de cet état à l'état correspondant au corpus vide¹ : on parlera de raffinements « impossibles »,
2. certaines transitions « bouclent » sur cet état² : on parlera de raffinements « connus »,³



Figure 5 – Exemple d'interactions homme-machine basées sur les notions de « raffinement connu », « inconnu » et « possible ».

¹ C'est le cas de *c* dans l'état $\{B\}$ (cf. Fig.4) : la requête *c* appliquée au corpus $\{B\}$ a pour résultat le corpus \emptyset .

² C'est le cas de *a* et de *b* dans l'état $\{B\}$: la requête *a* et la requête *b* ont pour résultat le corpus $\{B\}$.

3. pour les autres transitions, on parlera de raffinements « possibles ».

A la lumière de ces trois définitions, on peut construire, à partir d'un simple index de type documents/mots-clés (cf. Tab.1), un système de « navigation » entre corpus permettant d'assister l'utilisateur dans le raffinement de ses requêtes (cf. Fig.5).

Malgré l'intérêt que représente un tel système, on peut s'interroger sur son utilisation dans le contexte qui nous intéresse. En effet, l'archéologie est l'une des sciences ayant le vocabulaire le plus riche. Le nombre de mots-clés candidats pour l'indexation et la recherche risquerait donc de dépasser tant les capacités cognitives de l'utilisateur que calculatoires de l'ordinateur.⁴ Aussi, nous proposons d'adopter comme description des documents, non plus des descripteurs isolés, mais des descripteurs organisés dans un ordre partiel (cf. Fig.6) conformément à des relations explicites de spécialisation.⁵ Ainsi pour un état donné, il ne s'agit plus de donner une valeur *{connue, possible, impossible}* à tous les descripteurs mais uniquement aux spécialisations du descripteur « racine » (nommé *Vestige* dans la Figure 6) et aux spécialisations de ses spécialisations lorsque ces dernières sont « connues » (et ainsi de suite). La navigation entre corpus s'effectue alors à partir de la sélection du descripteur racine et par sélection successive de descripteurs « possibles » ou « désélection » de descripteurs précédemment sélectionnés. Suivons pas à pas le scénario présenté par la Figure 7 :

Etape 1. Le corpus global traite de « *vestiges typés* ». Les corpus plus spécialisés traitent de *vestiges datés*, de *vestiges de type épigraphique* ou de *vestiges de type instrumenta/mobilier* mais pas de *vestiges de type architectural* (cette description ne correspond en effet à aucun document du corpus considéré). Après sélection par l'utilisateur de *instrumenta/mobilier*, le système passe à l'étape 2.

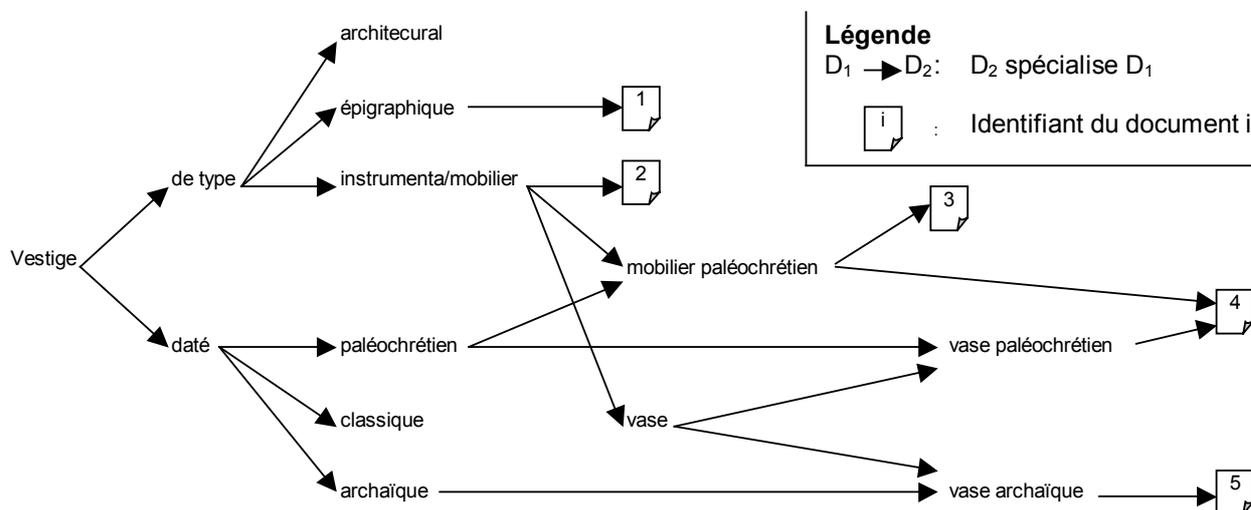


Figure 6 – Graphe orienté indexant cinq documents traitant de fouilles funéraires.

³ On précisera pour le lecteur attentif que la nature des raffinements à partir du corpus vide est non définie.

⁴ Dans un contexte interactif et dynamique.

⁵ Dans un réseau de concepts (cf. [SOWA92]) et non plus de descripteurs, notre relation de spécialisation entre le descripteur *vestige* et le descripteur *de type* deviendrait une relation « sorte de » entre le concept *vestige* et le concept *vestige typé*. De même, notre relation entre le descripteur *instrumenta/mobilier* et l'identifiant 2 deviendrait une relation « d'instanciation ».

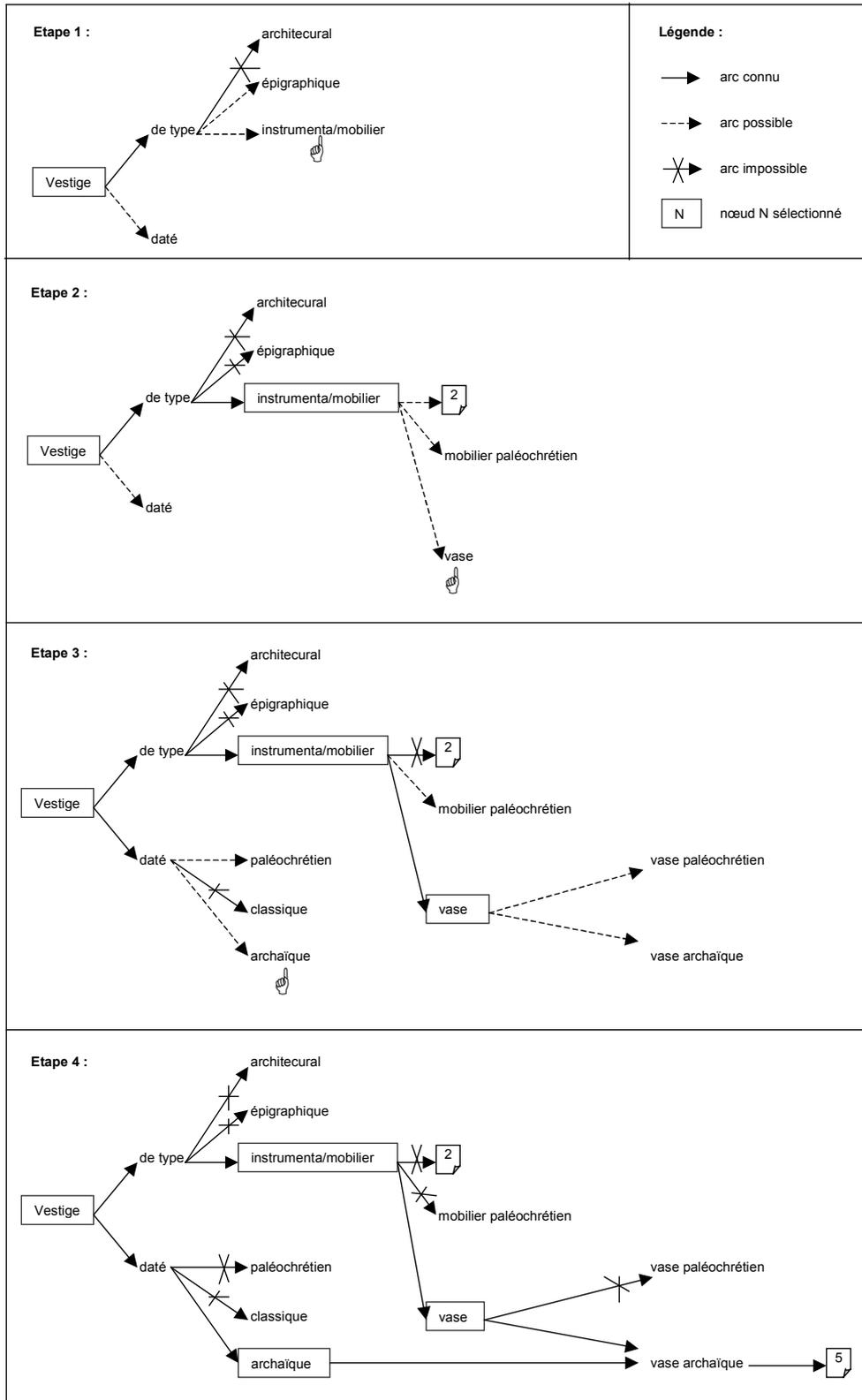


Figure 7 – Scénario en quatre étapes d’une navigation dans le graphe d’indexation de la Figure 6.

Etape 2. Le corpus sélectionné traite de *vestiges de type instrumenta/mobilier*. Cette description correspond exactement au document ayant 2 pour identifiant. Les corpus plus spécialisés traitent de *mobilier paléochrétien* ou de *vases*. Aucun ne traite de *vestiges de type architectural* ou *épigraphique*. Après sélection par l'utilisateur de *vase*, le système passe à l'étape 3.

Etape 3. Le corpus sélectionné traite de *vestiges de type vase (instrumenta/mobilier)* et de *vestiges datés*. On remarque que le fait qu'ils soient *datés* est inféré par l'ordinateur (tous les documents du corpus traitant de *vases* traitent de *vestiges datés*). Cette description correspond exactement au document ayant 2 pour identifiant. Les corpus plus spécialisés traitent de *mobilier paléochrétien*, de *vases paléochrétiens*, de *vases archaïques*, de *vestiges paléochrétiens* ou de *vestiges archaïques*. Aucun ne traite de *vestiges de type architectural* ou *épigraphique*, ni de *vestiges datés de l'époque classique*. Après sélection par l'utilisateur de *vase archaïque*, le système passe à l'étape 4.

Etape 4. Le corpus sélectionné traite de *vestiges de type vase (instrumenta/mobilier) datés de l'époque archaïque*. Ce corpus ne contient qu'un seul document : celui ayant 5 comme identifiant.

4.2 Indexation de documents

Tout comme la recherche de documents, le processus d'indexation d'un document dans un corpus peut être identifié comme étant un processus de *définition* d'un document, *définition* au sens de *distinction*. Il s'agit en effet de distinguer un unique document d'un ou, éventuellement, de plusieurs corpus.

Notre système de navigation parmi les corpus, permet à l'indexeur de se déplacer rapidement vers un corpus qui pourrait contenir le document à indexer mais dont les sous-corpus actuels ne le pourraient pas. Par exemple, pour ajouter un document traitant de *vases classiques* dans l'index représenté par la Figure 6, on pourra suivre les étapes 1 à 3 de la Figure 7. En effet, le document à indexer pourrait appartenir au corpus traitant des *vases* mais pas à ceux traitant des *vases paléochrétiens* ou *archaïques*.

L'indexeur peut alors rattacher l'identifiant de son document aux descripteurs connus, à certains descripteurs possibles, ou encore à des spécialisations de descripteurs connus qu'il créera pour l'occasion. L'opération peut se répéter afin d'associer l'identifiant du document ou les nouveaux descripteurs à d'autres corpus. Ainsi, toujours dans notre exemple, l'indexeur pourra rattacher un nouveau descripteur *vase classique* à *vase*, puis rattacher l'identifiant du document à *vase classique*, et enfin rattacher *vase classique* à *classique*.

Utiliser l'algorithme de navigation entre corpus non seulement pour la recherche mais aussi pour l'indexation, nous permet de faciliter la réutilisation des index précédents. Il s'agit en effet de favoriser les recoupements sans pour autant imposer un « vocabulaire contrôlé ».

4.3 Cas particulier des documents structurés

Nous proposons maintenant d'étendre aux documents structurés les notions introduites ci-dessus. Pour Franck Fourel (cf. [FOUR98]), la structure d'un document est organisée selon trois types de relations : celles de composition,⁶ celles de séquence⁷ et celles de référence.⁸ La question est alors

⁶ ex : un chapitre est composé de sous-chapitres

⁷ ex : un paragraphe suit un autre

⁸ ex : une phrase fait référence à une figure

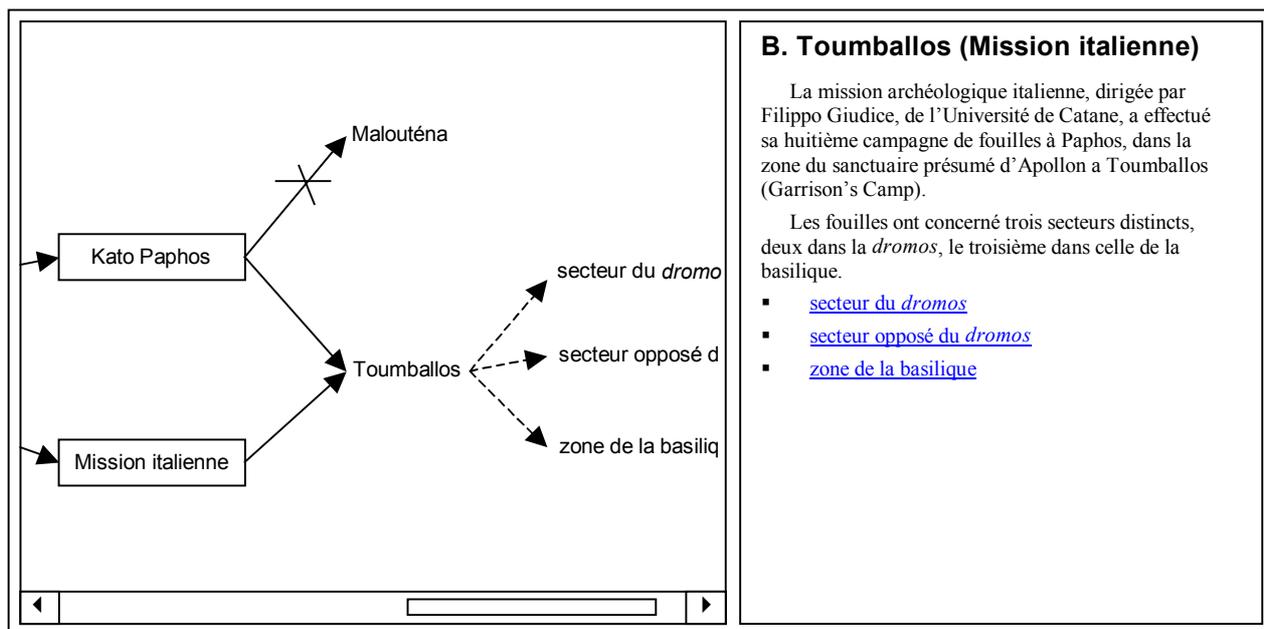


Figure 8 – Navigation conjointe dans le graphe d'indexation et le document structuré.

de savoir comment se propage l'attribut⁹ d'un élément structurel¹⁰ suivant ces relations. Pour reprendre son exemple, doit-on dire qu'un document est à l'état de brouillon lorsqu'au moins un chapitre est à l'état de brouillon, lorsqu'ils le sont tous, lorsqu'un certain pourcentage l'est ou enfin est-il impossible de déduire l'état du document à partir de l'état des chapitres ?

Tout d'abord, nous plaçant dans un contexte beaucoup plus limité, nous ne considérerons qu'un type de relation. En effet, d'une part, nous négligerons les relations de séquence qui ne nous paraissent pas indispensables pour la recherche proprement dite, et d'autre part nous regrouperons dans le même concept de « référence acyclique »¹¹ la composition et la référence. Nous serons donc amenés, dans le cas de références cycliques entre éléments structurels à considérer un élément virtuel faisant une référence acyclique à ces éléments.¹² Ensuite, dans un souci de simplicité, nous poserons que la description du tout est applicable à la partie et que la description de la partie n'est pas applicable au tout. Ainsi, il devient possible d'utiliser notre système non plus uniquement avec des documents mais aussi avec des parties de documents. Il est à noter que dans notre système, une même partie de document peut appartenir à la fois à plusieurs structures différentes. L'interface homme-machine est telle que tous les fragments de document correspondant à un « raffinement connu ou possible » (cf. définitions en 4.1.2 et 4.1.3) seront affichés automatiquement (ou sous forme iconique). Il devient alors possible d'afficher un index, une carte, le document entier, un titre, une introduction ou tout autre fragment de document multimédia pouvant être considéré comme une étape dans la navigation vers une partie plus spécialisée. Précisons que la navigation pourra se faire conjointement à l'aide du graphe d'indexation ou des hyperliens des fragments de document affichés.

⁹ ex : « auteur = Tartempion », « date_modification = 04/02/2000 »

¹⁰ ex : paragraphe, figure, chapitre, document...

¹¹ Les références acycliques sont définies par opposition aux références cycliques pour lesquelles il existe un chemin dont le départ et l'arrivée sont le même élément.

¹² ex : On pourra, par exemple, remplacer le fait que l'élément structurel traitant de *l'Artémision d'Ephèse* et celui traitant du *sanctuaire de Brauron* se fassent mutuellement référence par le fait qu'un nouvel élément structurel traitant des *sanctuaires d'Artémis* leur fasse référence.

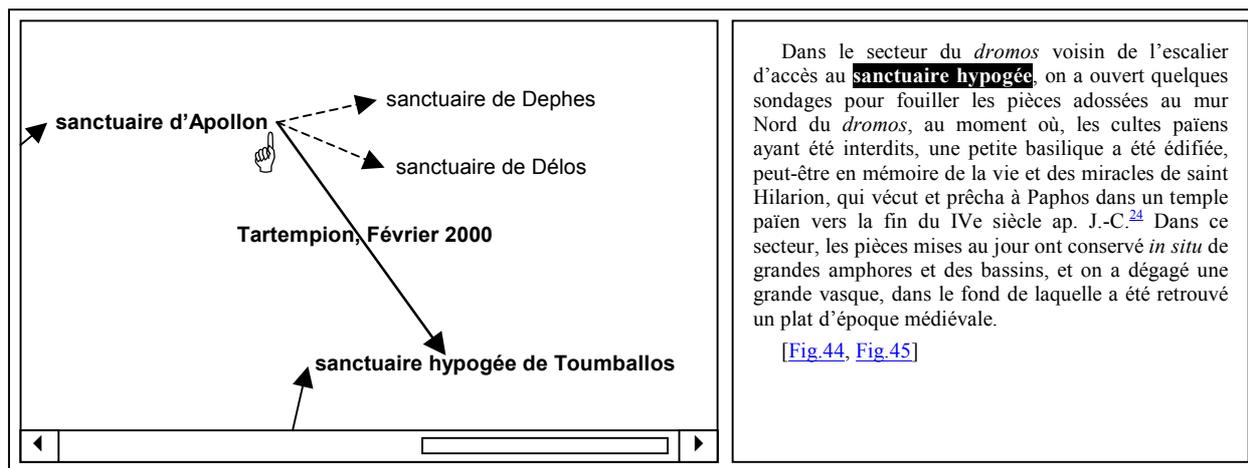


Figure 9 – Annotation dans un document par un lecteur.

Dans l'exemple de la Figure 8, un fragment de document a été affiché à la suite d'une recherche multicritère dans le graphe d'indexation (le lieu *Kato Paphos* et l'entreprise de fouille *Mission italienne* ont été sélectionnés). D'autres fragments de documents, correspondant à des secteurs géographiques plus précis (*secteur du dromos*, etc.), peuvent être atteints soit à l'aide des hyperliens contenus dans le fragment de document soit par sélection des descripteurs correspondants.

4.4 Annotation et collaboration

Une annotation dans notre contexte est l'activité d'un utilisateur consistant à relier (par une relation de généralisation) un descripteur ou un identifiant (par exemple celui d'un mot dans un document) à un autre descripteur. Il est à noter que le descripteur ou l'identifiant annoté aura pu être créé par ce même utilisateur, par un autre ou par l'indexeur.

Dans l'exemple de la Figure 9, Monsieur Tartempion, après avoir souligné *sanctuaire hypogée* dans le texte, et l'avoir identifié par *sanctuaire hypogée de Toumballos*, a rattaché (en février 2000) cet identifiant au descripteur préexistant *sanctuaire d'Apollon*.

Martin Röscheisen et son équipe (cf. [ROSC95]), se sont intéressés à la question des annotations dans le contexte d'une communauté « en-ligne ». Dans leur modèle, sont gérés les informations portant sur la création des annotations, ainsi que les droits de lecture et de modification. La répartition des annotations sur le réseau est fonction des droits qui y sont attachés, afin d'obtenir des serveurs spécialisés pour des groupes d'utilisateurs.

De même, dans notre système, afin d'éviter de perdre la « sémantique situationnelle »¹³ de l'annotation, chaque arc du graphe est étiqueté par son auteur et sa date de création. D'autre part, il est possible, de définir des accès publics (par exemple pour des travaux publiés), privés (pour des travaux non-publiés) ou réservés à un groupe (par exemple pour les étudiants d'un cours ou les membres d'une équipe de recherche).

En ce qui concerne la répartition des données et leur réplique, nous envisageons de généraliser l'approche en l'appliquant non plus seulement aux droits d'accès mais à la consultation effective

¹³ Cette théorie linguistique de la sémantique veut que le sens d'un mot soit donné en partie par la *situation* dans laquelle il a été prononcé (personne, instant, lieu, etc.), à ne pas confondre avec une autre théorie qui veut qu'il soit donné par l'ensemble de ses *contextes* d'utilisation (c'est à dire des mots qui l'entourent).

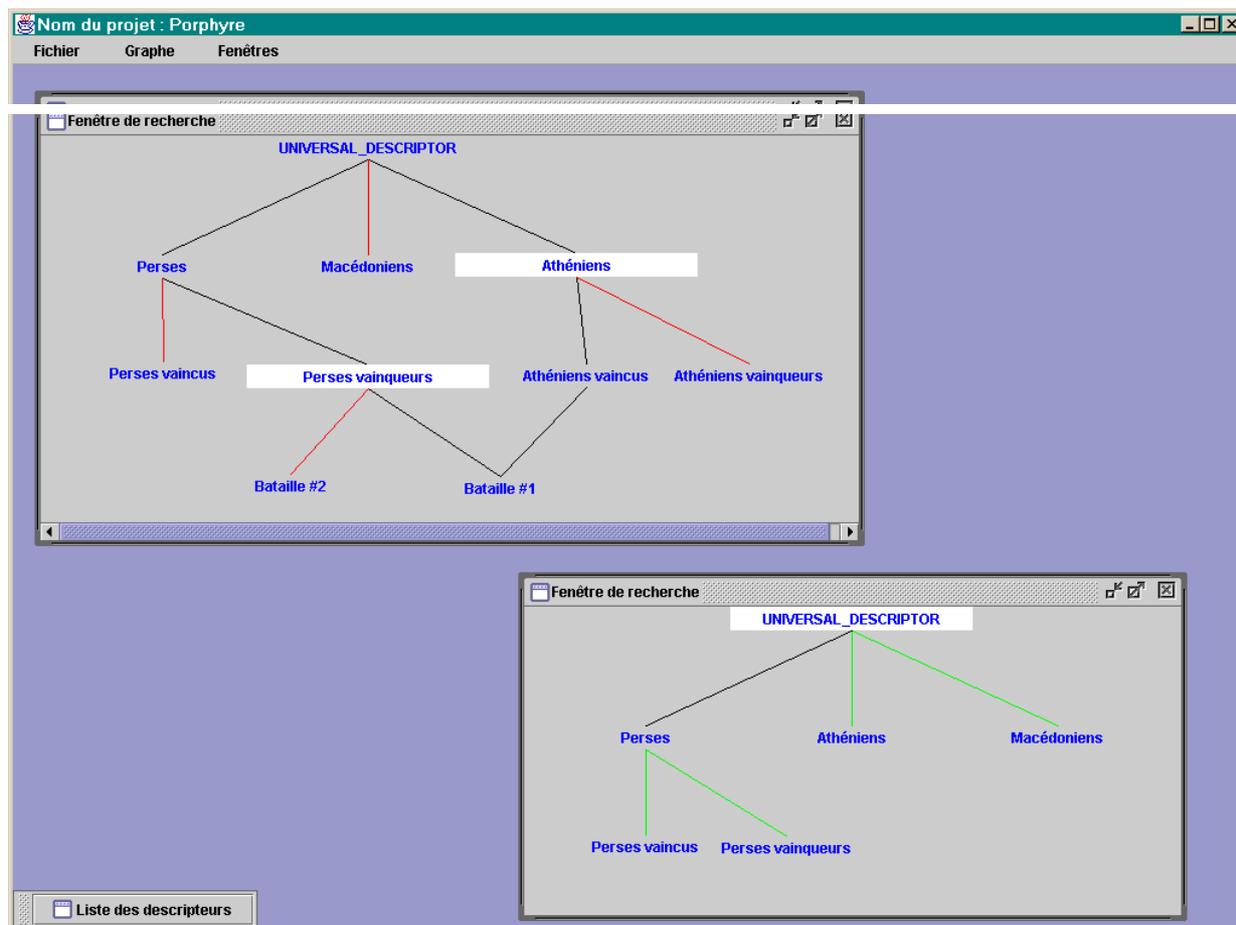


Figure 10 – Copie d'écran du prototype en Java du module de navigation.^{16 17}

des annotations ainsi qu'à la notion de proximité¹⁴ des utilisateurs. Il s'agirait en effet d'utiliser pleinement la métaphore de la diffusion des savoirs au sein de la communauté scientifique.¹⁵

5 Bilan et perspectives

Nous nous sommes donc intéressés aux tâches documentaires de chercheurs (dans notre cas, des chercheurs en archéologie) face à un corpus de « publications savantes ». Tout d'abord, nous avons intégré la recherche de documents dans une activité plus large, plus complexe, impliquant d'avantage l'utilisateur et comprenant entre autres l'enrichissement de la description des documents et le partage de cette description avec d'autres utilisateurs. Ensuite nous avons introduit une approche de la recherche de documents basée sur la navigation entre corpus, mettant ainsi en avant l'utilisateur, le corpus de documents et leurs relations (interactivité, aspect exploratoire...). Enfin nous avons défini comment un outil filtrant un graphe d'indexation en

¹⁴ On se rapproche ici de la notion de « proxy ».

¹⁵ Les interactions entre chercheurs sont dépendantes des thèmes de recherche mais aussi de leur proximité géographique ou institutionnelle.

¹⁶ Le fait de différencier l'état des arcs par des couleurs et non des formes ne résulte pas de choix ergonomiques pour le long terme mais seulement de choix de simplification pour la version initiale du prototype.

¹⁷ La direction globale du graphe acyclique a été revue depuis le développement du prototype (non plus de haut en bas mais de gauche à droite).

fonction de la sélection de descripteurs permettait de rechercher et d'indexer des documents simples ou structurés, ainsi que de les annoter et d'échanger ces annotations.

Le travail présenté ici s'inscrit dans le cadre du projet de « mise en ligne » de la *Chronique des fouilles*, projet lié à la thèse d'Aurélien Béné¹⁸ et qui a fait l'objet de recherches sur l'application de la théorie « situationnelle » de la sémantique en indexation (cf. [BENE99]). A l'heure actuelle, nous disposons d'un prototype en Java (cf. Fig.10) qui nous a permis de mettre en évidence la validité de l'algorithme de filtre de graphes acycliques ainsi que la faisabilité d'une interface homme-machine basée dessus.

Dans les mois à venir, nous prévoyons d'éprouver notre prototype, tant au niveau algorithmique avec de gros volumes de données, qu'au niveau interface auprès des utilisateurs (pour cela nous devons faire évoluer le prototype vers une version client-serveur).

Ainsi, dans les prochaines années, nous pourrions offrir aux chercheurs un système leur permettant de prendre des « notes de lecture », de les partager, et de « naviguer » à travers le réseau qu'elles constituent ; le pari étant, tel que le souhaitait Vannevar Bush (cf. [BUSH45]), d'assister le lecteur de publications savantes en le soulageant des aspects répétitifs de son activité et en le laissant se concentrer sur les aspects créatifs, intuitifs et à haut niveau d'abstraction.

Références :

- [BENE99] : Aurélien Béné, Sylvie Calabretto et Jean-Marie Pinon. Indexation « sémantique » de documents archéologiques. In *Actes du deuxième colloque du chapitre français de l'ISKO, « L'indexation à l'heure d'Internet »*, 1999.
- [BUSH45] : Vannevar Bush. As we may think. *The Atlantic Monthly*. July 1945.
- [CARP94] : Claudio Carpineto et Giovanni Romano. Dynamically bounding browsable retrieval spaces: an application to Galois lattices. In *RIAO'94 Conference Proceedings, "Intelligent Multimedia Information Retrieval Systems and Management"*, 1994.
- [CHIA99] : Yves Chiaramella. Approches et modèles en recherche d'informations. In *XVIIe congrès INFORSID, Journée de cours : « Recherche d'informations »*, 1999.
- [FOUR98] : Franck Fourel. *Modélisation, indexation et recherche de documents structurés*, chapitre 6 : Indexation structurelle. Thèse de doctorat en informatique, Université de Grenoble 1, 1998.
- [LAI97] : Michel Lai. *UML La notation unifiée de modélisation objet : Applications en Java*. Masson, Paris, 1997.
- [OHAR98] : Kenton O'Hara, Fiona Smith, William Newman et Abigail Sellen. Student readers' use of library documents: Implications for library technologies. In *CHI'98 Conference proceedings on Human factors in computing systems*, 1998.
- [ROSC95] : Martin Röscheisen, Christian Mogensen et Terry Winograd. Beyond browsing: Shared comments, soaps, trails, and on-line communities. In *The Third International World Wide Web Conference, "Technology, Tools and Applications"*, 1995.

¹⁸ Thèse en cours sur la modélisation de la documentation de consultation pour experts et son application à l'archéologie, sous la cotutelle du Laboratoire d'Ingénierie des Systèmes d'Information et de l'Ecole française d'Athènes.

- [SALT68] : Gerard Salton. *Automatic Information Organization and Retrieval*, chapitre "Retrieval Models". Computer Sciences series. McGraw-Hill Inc., USA, 1968.
- [SOWA92] : John F. Sowa. Semantic Networks. In *Encyclopaedia of Artificial Intelligence*, Wiley, New York, 1987 ; seconde édition revue et corrigée, 1992.
- [TOCH94] : Klaus Tochtermann. A first step toward communication in virtual libraries. In *Digital Libraries '94: Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries*, 1994.