
Assistance à l'interprétation dans les bibliothèques numériques pour les sciences historiques

Aurélien Bénel* — **Sylvie Calabretto**** — **Bruno Helly***** —
Andrea Iacovella****

* *Equipe Tech-CICO, Laboratoire ISTIT (CNRS-UTT)*
Université de Technologie de Troyes,
12 rue Marie Curie, BP 2060, 10010 Troyes CEDEX
aurelien.benel@utt.fr

** *Laboratoire LIRIS (CNRS-INSA-Lyon1-Lyon2-ECL)*
Institut National des Sciences Appliquées de Lyon,
20 avenue Albert Einstein, 69621 Villeurbanne CEDEX
sylvie.calabretto@insa-lyon.fr

*** *Fédération de Recherche MOM (CNRS)*
5/7 rue Raulin, 69007 LYON
bruno.helly@mom.fr

**** *Centre d'Édition Numérique Scientifique (CNRS)*
École Normale Supérieure Lettres & Sciences Humaines,
15 Parvis René Descartes BP 7000, 69342 Lyon Cedex 07
andrea.iacovella@ens-lsh.fr

RÉSUMÉ. Cet article s'intéresse à la question de l'interprétation dans les bibliothèques numérique, et à son assistance à l'aide d'un outil informatique. Dans ce but, nous présentons un « espace de manipulation de traces » conçu au sein d'un réseau transdisciplinaire. Au cours de l'article, trois modes de construction du sens sont illustrés par des exemples puis discutés par les intervenants des différentes disciplines (archéologie, histoire, historiographie, informatique).

ABSTRACT. How to assist digital libraries readers in interpreting documents? The proposition of our transdisciplinary team is to offer them a "digital trace handling space". Throughout this article we focus on three complementary ways to "make sense" with documents. Each one is illustrated with an example and then discussed by the team members from their own disciplinary point of view (archaeology, history, historiography, computer science).

MOTS-CLÉS : transdisciplinarité, bibliothèques numériques, points de vue, travail coopératif, annotation, modélisation des connaissances, hypermédia.

KEYWORDS: transdisciplinarity, digital libraries, viewpoints, cooperative work, annotation, knowledge modelling, hypermedia.

1. Introduction

Le point de départ de cette étude réside dans un questionnement théorique sur la construction du sens dans les sciences historiques. En particulier, nous avons eu l'occasion à plusieurs reprises d'identifier comme une impasse méthodologique le courant de l'archéologie défendant la définition normative d'une typologie des vestiges (ou l'utilisation de thesaurus pour l'interrogation de « banques de données »).

À la différence du spécialiste des Sciences de l'Information, occupé à affecter des significations non ambivalentes et non ambiguës en regard des corpus documentaires, l'historien manipule un objet qui déborde de sens de toutes parts. Dans un index, un catalogue, un titre d'ouvrage ou dans un corps de texte, le mot « Auschwitz » ne se laisse pas résumer à de l'information. Écrit en toutes lettres ou représenté par son équivalent binaire, Auschwitz reste Auschwitz dans toute son horreur.

L'émergence au sein de la communauté des historiens d'effrayants non-sens, comme le négationnisme, montre que la dynamique interprétative est en panne, coincée entre une subjectivité débridée et une pseudo-objectivité. Ainsi, le but scientifique de l'équipe est d'explorer et si possible de renouveler les méthodes interprétatives dans les sciences historiques.

Or, l'interprétation fait partie de ces phénomènes humains qui, selon Michel de Certeau, sont découpés, analysés, expliqués par différentes disciplines scientifiques en fonction de leurs postulats et de leurs méthodes. Dès lors, le chercheur intéressé par cette question se trouverait acculé entre un discours schizophrénique et un « méta-discours » censé dépasser les disciplines (De Certeau, 1987). La posture transdisciplinaire dispose donc d'une marge de manœuvre bien étroite et demande à tout moment d'inventer de nouvelles méthodes en particulier pour l'expérimentation et la publication.

Pour l'expérimentation, nous tentons de mettre en place des plateformes d'expertise. Chacune est constituée autour d'une communauté d'experts et de leur corpus documentaire (intégré à un prototype informatique). Elle est le lieu de mise à l'épreuve du prototype et par là du modèle théorique élaboré dans le dialogue transdisciplinaire.

En ce qui concerne la publication des résultats, notons qu'elle nécessite de nouvelles formes d'écriture. Dans cet article, par exemple, nous partirons de la présentation de l'objet transdisciplinaire : non pas le modèle informatique, qui finalement ne fait que définir la structure d'un espace, mais son appropriation par les usages. Ensuite, nous ouvrirons, autour de chaque thème abordé, une discussion entre des membres de l'équipe issus de différentes disciplines.

2. Structure de l'espace de manipulation de traces

Afin d'instrumenter un travail humain d'interprétation, nous proposons un espace informatique de manipulation de traces. Tout comme la trace d'un animal dans la neige, il s'agit d'une forme qui demeure en l'absence de son contenu. Autrement dit, nous stockerons un symbole dénué de signification pour l'ordinateur mais qui n'en restera pas moins re-sémiotisable par l'être humain.

2.1. Indices pour la compréhension des traces

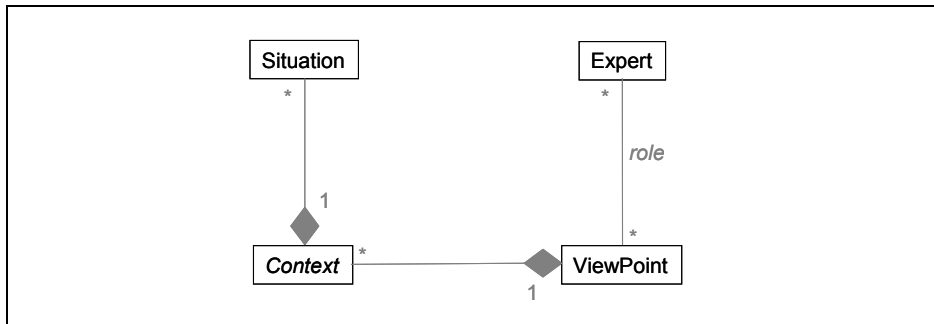


Figure 1. Indices pour la compréhension des traces (modèle de classe UML)

CONTEXTE – Une trace sera d'autant mieux comprise qu'elle sera en relation avec d'autres. Chacune de ces relations sera appelée « contexte ».

POINT DE VUE – Ces mises en contexte se rassemblent pour former le point de vue d'une communauté, communauté dans laquelle les rôles peuvent être très différents.

EXPERT – L'être humain considéré n'est pas un « utilisateur » dont les tâches se borneraient à de la consultation ou à du remplissage de formulaires, mais un expert. En tant que spécialiste, il structure lui-même son objet d'étude et son activité.

SITUATION – A chaque contexte est associé son historique. Les événements de cet historique (transition entre les états : brouillon, publié, obsolète et pilon) sont appelés des situations. Chacune répond aux questions « quoi » (création/suppression/archivage), « quand » (horodatage), et « qui » (expert – éventuellement externe au système).

2.2. Différentes manières de contextualiser les traces

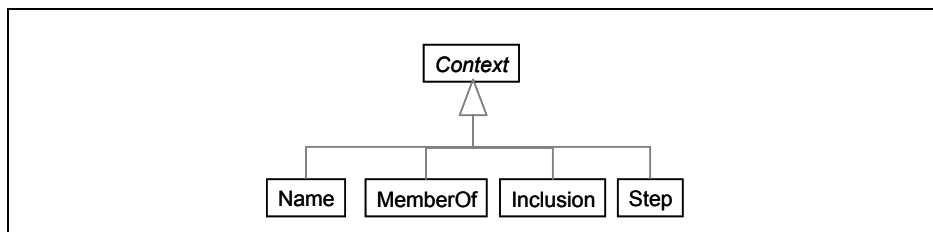


Figure 2. Les différents types de contexte (modèle de classe UML)

Le modèle propose différentes manières de contextualiser des traces (dénomination, regroupement, inclusion, étape de lecture). En tant que contexte, chacune est associée à un point de vue, ainsi qu'à l'historique des situations qui le concernent.

DÉNOMINATION – A première vue, donner un nom à une trace (collection, parcours...) reviendrait à affecter une valeur à un attribut. Cependant, ce nom pouvant varier (en particulier dans le temps) et cette variation étant significative pour l'expert, il nous a paru nécessaire de considérer cette dénomination comme un contexte particulier.

REGROUPEMENT – L'une des traces les plus importantes du système est l'objet documentaire (source ou fragment de texte ou d'image). Le fait de placer un objet documentaire dans une collection sera appelé « regroupement ». Notons qu'un même objet documentaire pourra être regroupé dans différentes collections, ces collections pouvant même appartenir à différents points de vue¹.

INCLUSION – Les collections sont organisées entre elles selon une relation d'inclusion. La collection englobante regroupe alors virtuellement tous les objets documentaires regroupés par la collection englobée (et ainsi de suite). D'un point de vue formel, les inclusions (actives à un moment donné²) forment un graphe orienté acyclique³.

¹ En effet, les objets documentaires sont les seuls points de jonction entre points de vue. Ces derniers peuvent être développés de manière indépendante et même stockés sur des serveurs distants les uns des autres.

² Si l'on considère le *graphe des inclusions* de manière diachronique, le graphe orienté n'est plus acyclique mais quelconque.

³ Les graphes étaient beaucoup plus contraints dans les versions antérieures de *Porphyry*. Dans un point de vue il ne pouvait y avoir qu'un seul graphe connexe. De plus, les nœuds du graphe étaient typés de manière forte : « facettes » pour les

ÉTAPE DE LECTURE – On appelle étape d'un parcours de lecture, le fait de lire un objet documentaire après un autre. Pour un parcours donné, l'enchaînement de ces étapes (actives à un moment donné⁴) est tel que le parcours de lecture est de type "visite guidée" : on ne repasse jamais deux fois au même endroit. Plusieurs parcours pouvant passer par le même objet documentaire, il est par contre possible en un objet documentaire de bifurquer d'un parcours vers un autre.

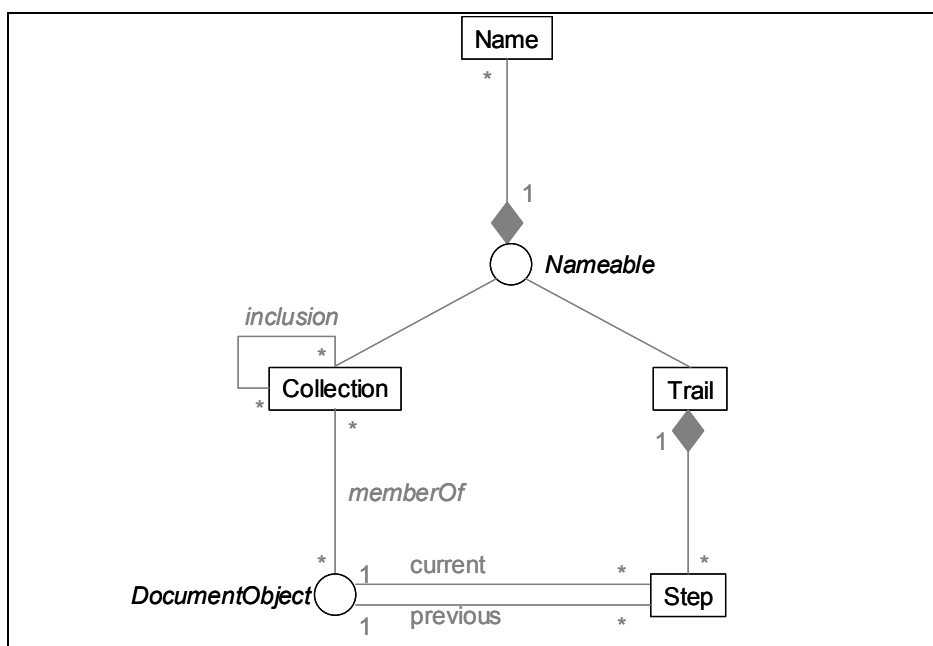


Figure 3. Règles d'utilisation des différents types de contexte (modèle de classe UML)

sources, « identifiants » pour les puits, « descripteurs ordinaires » pour les autres. Enfin, chaque graphe connexe avait une et une seule facette ; et les identifiants ne pouvaient faire référence qu'à un seul objet documentaire. A l'usage, il appert que ce typage fort et ces contraintes excessives s'avèrent incompatibles avec le caractère dynamique de la modélisation.

⁴ Si l'on considère le *graphe des étapes* (pour un parcours donné) de manière diachronique, le graphe orienté ne sera plus une liste mais sera quelconque.

3. Appropriation de l'espace par les usages

3.1. Intertextualité

3.1.1. Etude de cas

Le premier usage considéré est celui d'un archéologue face au rapport de fouille d'une nécropole, un rapport rédigé par d'autres mains (Orsi et Cavallari) en une autre époque (fin XIX^e s.). Le rapport extrêmement bien structuré indique pour chacune des sépultures, le type de tombe, les restes du(des) défunt(s) ainsi que le mobilier retrouvé.

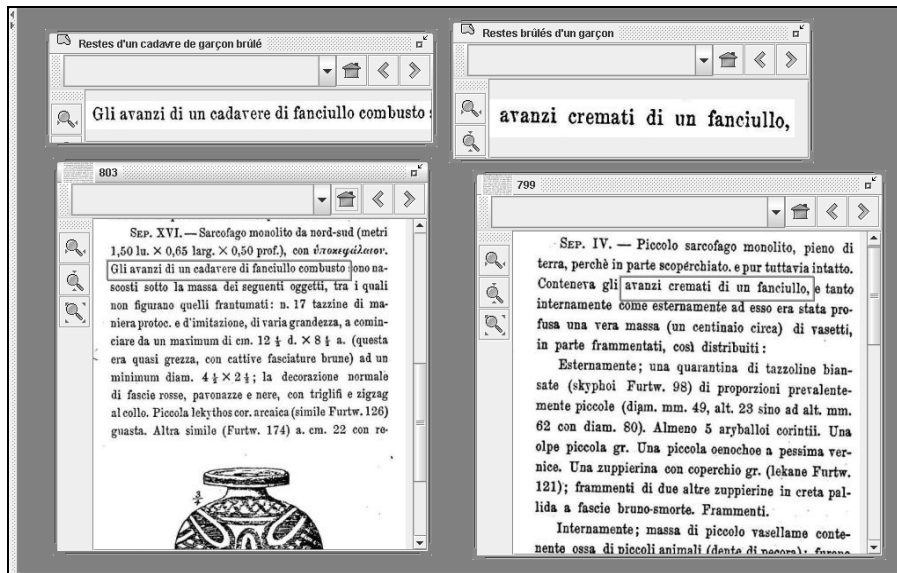


Figure 4. Marquage par le lecteur d'une analogie entre deux documents sources (capture d'écran du système Porphyry)

Dans le rapport, l'attention de l'expert est attirée par une mention surprenante : « restes d'un cadavre de garçon brûlé ». En effet, si « brûlé » signifie qu'il est question d'une crémation, que peut signifier « restes de cadavre » ? De plus, toujours dans le cadre d'une crémation, et donc en absence de squelette, comment le fouilleur a-t-il pu déterminer l'âge (jeune) et le sexe (masculin) du « garçon » ?

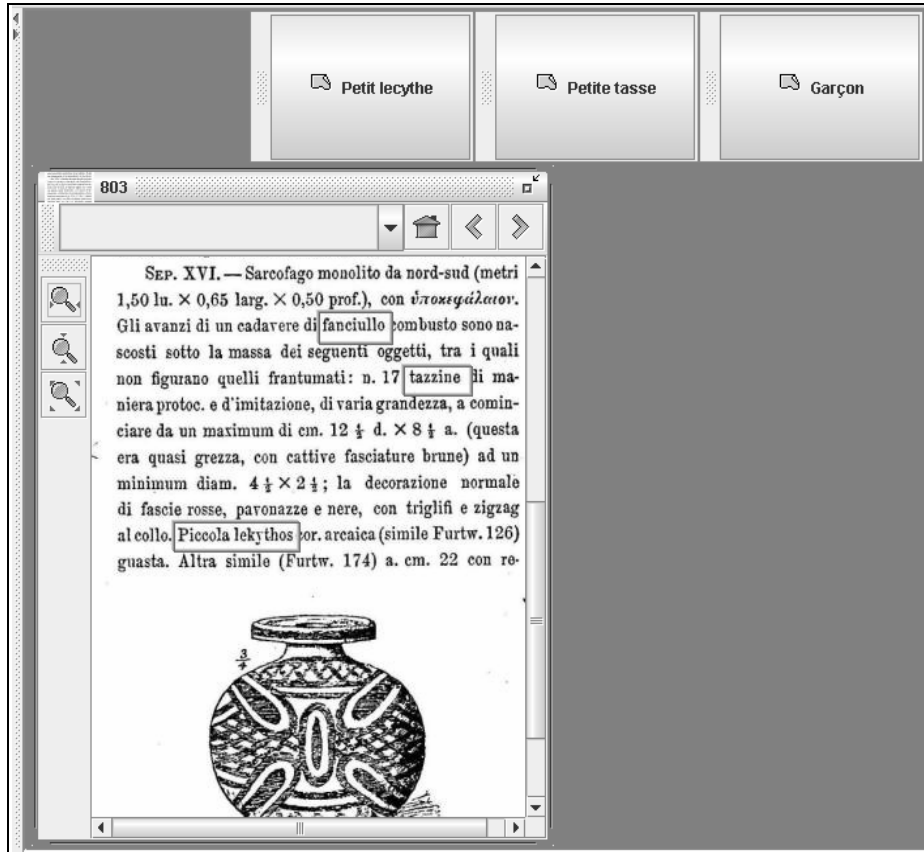


Figure 5. Marquage par le lecteur d'analogies au sein d'un même document source (capture d'écran du système Porphyry)

Cette « bizarrerie » commence à prendre sens, lorsque l'expert découvre qu'elle est récurrente dans une autre sépulture où il est question de « restes brûlés d'un garçon ». Ainsi, il ne s'agit probablement pas d'une erreur, mais du fruit de l'expertise de l'auteur, dont les étapes se cachent peut-être « entre les lignes ».

Une étude approfondie de la description de la sépulture montre que le thème de la petite taille présent dans le mot « garçon » est également présente dans « petite tasse » ou encore « petit lecythe » (sorte de vase). L'archéologue peut alors faire l'hypothèse qu'en l'absence de squelette, le fouilleur a déterminé l'âge du défunt à partir de l'étude du matériel de la sépulture.

3.1.2. *Discussion interdisciplinaire*

Andrea Iacovella (historiographie, archéologie)

Cette étude de cas illustre les travaux que nous menons depuis plus de quinze ans à la croisée de l'archéologie, de l'historiographie (histoire de l'histoire) et des sciences cognitives. Ces travaux visent à analyser le discours des archéologues afin d'en extraire de nouvelles conclusions archéologiques. En quelque sorte, il s'agit de refaire, virtuellement, une fouille effectuée au siècle dernier.

L'enjeu est de taille, puisque l'archéologie a pour particularité de détruire son objet d'étude. En effet, même si les musées et les apothèques sont remplies de vestiges, leur contexte est à jamais détruit par l'opération même de la fouille. Mais, alors que la couche archéologique est détruite pour passer à celle du dessous, une strate documentaire est ajoutée au corpus. L'unique fouille archéologique que nous laissons aux générations futures est donc la « fouille de documents⁵ ».

Qu'il s'agisse d'établir des relations internes au discours de l'archéologue sur une fouille, de relations avec d'autres rapports de fouille (en faisant varier les archéologues, les sites, les campagnes de fouille), ou encore de relations avec des documents d'archive, cette fouille virtuelle se fait en tissant des relations de sens entre parties de documents.

Aurélien Bénéel (informatique)

Existe-t-il des méthodes qui soient communes aux disciplines (pourtant si différentes) regroupées sous l'appellation « Sciences humaines » ? Des travaux comme ceux de Paul Ricœur tendent à rattacher les Sciences humaines à la tradition herméneutique, science de l'interprétation issue de l'étude des textes sacrés (Ricœur, 1986). L'un de ses principes est l'intertextualité, c'est-à-dire l'idée que deux textes (au sens large) lus en parallèle sont d'avantage porteurs de sens que lus séparément, comme s'ils sélectionnaient du sens l'un dans l'autre (cf. François Rastier).

L'idée d'instrumenter sur ordinateur cette « sélection du sens » n'est pas nouvelle. Le projet Xanadu® est de ceux-là. Il s'agit du projet que poursuit, depuis les années 1960, Ted Nelson (philosophe et sociologue), inventeur des termes « hypertexte » et « hypermédia ». Le principe consiste à lire deux textes en parallèle en exhibant les relations d'analogie indiquées par le lecteur ainsi que les relations de citation indiquées par l'auteur (Nelson, 1999).

Plus largement, l'idée de constitution dynamique d'un corpus comme instrumentation du travail des experts rejoint les travaux de Jean Charlet et de son équipe autour des « dossiers numériques ». Comme ils l'ont analysé pour le domaine

⁵ Notons que si l'objet d'étude n'est plus seulement l'objet mais aussi le document, ceci repose le problème de démarcation entre l'histoire et l'archéologie.

médical, le dossier numérique permet par sa forme documentaire, son aspect contextuel, et la pluralité des auteurs et des lecteurs qui le manipulent de rendre compte d'un domaine où les objets (voire les sujets) d'étude sont complexes et où les modèles évoluent (Charlet et al., 1999).

Bruno Helly (archéologie, histoire, géographie)

Ce qui se pratique habituellement dans nos disciplines en matière de technologies de l'information, vise à stocker la description des vestiges dans des bases des données, la bibliographie dans des systèmes documentaires, et les textes, les photographies et les dessins dans le système de fichier. En réalité, un système d'information scientifique devrait intégrer un continuum des informations. Mais quel serait alors le statut de ces informations ?

Tout d'abord, une étude sur la « description factuelle des objets » montre qu'elle est le fruit d'une perception. Cette perception est sélective : c'est le problème posé ou le choix de l'objet d'étude qui conditionne la prise d'information. Ensuite, contrairement à une opinion bien ancrée dans l'esprit des archéologues, la perception d'un objet ne donne pas d'emblée une datation ou un classement typologique. Ces derniers sont des résultats, les produits d'un raisonnement complexe. Les archéologues considèrent que ce sont les objets qui parlent, alors que ce sont eux qui parlent sur les objets.

Même l'identification de l'objet d'étude n'a pas de permanence : les numéros et les catalogues d'inventaires changent, disparaissent ou au contraire se multiplient. De ce point de vue, seule une image de l'objet traité est à même d'assurer une bonne reconnaissance de l'objet d'étude. L'image elle-même dépend non seulement de la culture de l'observateur, mais aussi de la culture collective de la société pour laquelle on « fait l'histoire ». Une photographie est elle aussi « une vision du monde ».

En ce qui concerne les textes, on distingue trop souvent les textes issus d'observateurs directs (des événements passés) des textes de « seconde main », et ceci avec une connotation péjorative évidente pour les seconds. Or, en ce qui concerne la description d'un objet archéologique, qu'il s'agisse d'une déclaration formulée par un utilisateur du passé, qu'il s'agisse du catalogue d'un « antiquaire » du XVIII^e siècle, de l'article scientifique d'un de nos contemporains, ou encore d'un de nos propres post-it®, chacun est une source (discours daté et situé) qui par accumulation forme la bibliographie de l'objet. Nous défendons en quelque sorte une approche linguistique et philologique de l'archéologie, dans laquelle les objets n'ont pas de propriétés par eux-mêmes, sinon quand celles-ci sont déclarées dans des sources.

Sylvie Calabretto (informatique)

La philologie s'occupe de textes écrits dans des langues anciennes : l'ancien français, l'ancien occitan/provençal, l'italien et l'espagnol médiévaux, etc. Pour le philologue, il s'agit dans un premier temps de chercher à comprendre ces écrits, du point de vue littéral d'abord, puis dans leur signification profonde. Pour un texte original, il existe de 3 à 800 copies (appelées « témoin » en philologie), chaque témoin présentant des variantes de textes. L'objectif du philologue est de proposer son interprétation du texte en confrontant les différentes variantes (des mots) du texte. Le domaine de la philologie illustre ainsi parfaitement l'intertextualité puisqu'il s'agit de confronter différentes versions d'un texte original.

Dans le cadre du projet BAMBI (Better Access to Manuscripts and Browsing of Images) du programme LIBRARIES, nous avons développé une station de travail destinée à aider le philologue dans son activité. Les fonctions philologiques implémentées concernent la transcription et l'étude des variantes de textes (via l'apparat critique).

La transcription d'un manuscrit est un processus qui vise à noter la prononciation d'une langue donnée à l'aide d'un système de signes d'une langue de conversion. Le système BAMBI permet d'éditer et de visualiser la transcription dans une fenêtre adjacente à celle du document source. De plus, le système permet de faire correspondre automatiquement chaque mot de la transcription avec la portion de l'image source dans lequel le mot est trouvé. D'autre part, il est possible d'indexer la transcription et d'effectuer des recherches de mots ou de variantes.

L'apparat critique représente une zone dans laquelle le philologue réfère les variantes des mots du manuscrit et sa propre interprétation (exemple : *enssens*, *eixens*, *yssent* ou *eychins* sont des variantes de texte). Ainsi le philologue peut proposer sa lecture du texte et son interprétation.

3.2. Intersubjectivité

3.2.1. Etude de cas

La seconde étude de cas porte sur les « Collections de l'Ecole française d'Athènes en ligne⁶ ». L'Ecole française d'Athènes est un établissement public de recherche dans les disciplines se rapportant au monde grec. En tant que maison d'édition, elle publie une douzaine de collections (séries de monographies et revue). Depuis 1877, ce sont près de 570 volumes, soit 250.000 pages, qui ont été édités par l'Ecole. Si nous avons choisi l'exemple de ce corpus, c'est qu'au sein même de l'Ecole, différents acteurs vont l'organiser en fonction de leur spécialité.

⁶ Corpus disponible sur le Web à l'adresse : <http://cefael.efa.gr>

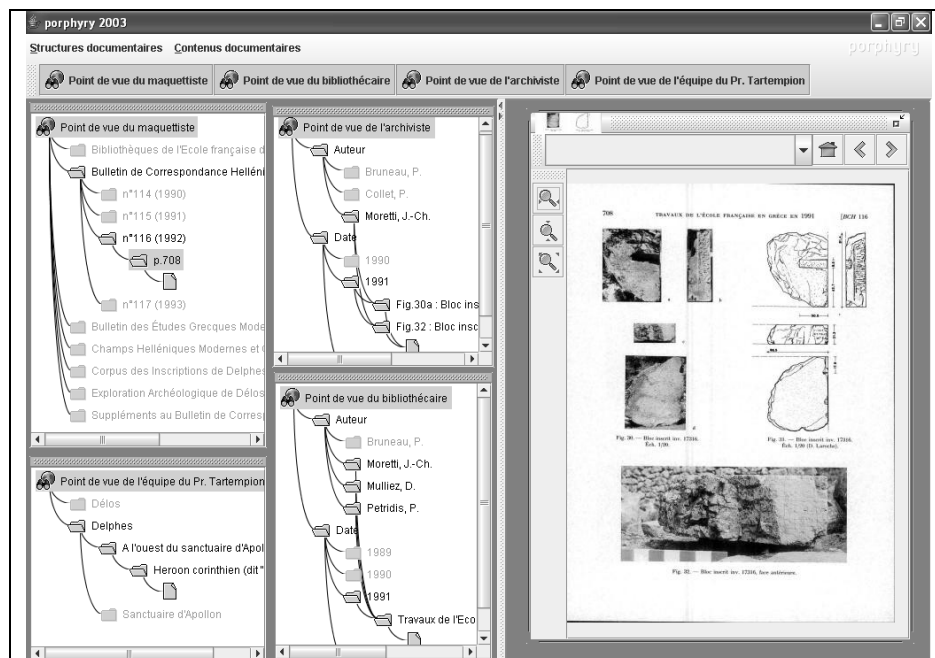


Figure 6. Feuilletage conjoint du corpus à travers plusieurs structures (capture d'écran du système Porphyry)

La première structuration du corpus est donnée par le maquettiste. Chaque page est désignée sans ambiguïté par le triplet « Collection/Volume/Folio ». Notons que c'est cette structure qui jusqu'à aujourd'hui est utilisée dans les bibliographies pour référencer un extrait.

La seconde est donnée par le bibliothécaire afin de faciliter l'accès au corpus pour les chercheurs. Cette structure identifie au sein des volumes des unités autonomes (articles de recherche, rapports, chroniques...) ayant pour attributs un titre, une date et un ou plusieurs auteurs.

L'Ecole dispose d'un fond de près de 500.000 photographies et plans datant de la fin du XIX^e siècle à nos jours. Ce fond comprenant entre autres les figures publiées dans les collections, on peut considérer que la structure du fond constitue une troisième structure pour la collection. Ainsi, à chacune des figures présentes dans la page sera associée la référence et la description (auteur, date de prise de vue) données par l'archiviste.

Enfin la quatrième structure ouvre sur beaucoup d'autres : celles des lecteurs. Dans notre exemple, il s'agit de l'analyse, par une équipe de recherche, de la

bibliographie concernant un site en fonction de la position spatiale de chacun des vestiges décrits⁷.

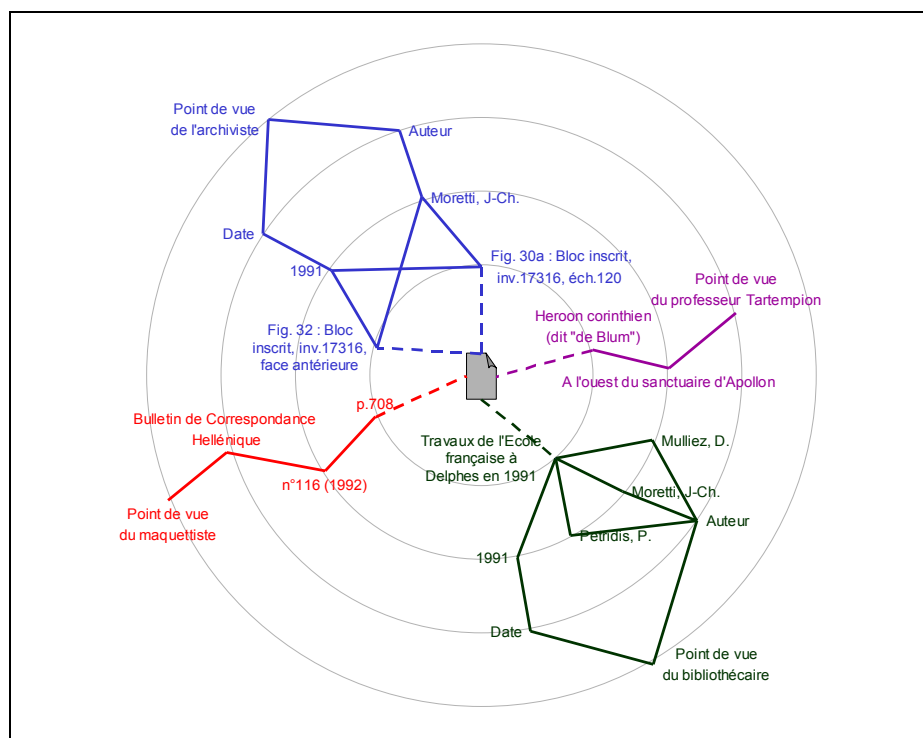


Figure 7. Visualisation du réseau centrée sur l'objet documentaire (maquette)

En fin de compte, chaque page du corpus vient à être décrite par un réseau complexe de points de vue qui se croisent.

3.2.2. Discussion interdisciplinaire

Aurélien Bénel (informatique)

D'un point de vue informatique, le thème de la construction du sens connote avec celui du Web sémantique, c'est-à-dire avec un approfondissement des rapports qu'entretiennent, sur un support partagé, des documents et des modèles de connaissance. Parmi les travaux de cette thématique, on peut distinguer trois manières d'envisager ce rapport :

⁷ Cet exemple fictif s'inspire des travaux actuels de l'équipe de Roland Etienne sur le sanctuaire de Délos.

- l'intelligence artificielle (« des machines comprennent des humains »),
- l'interopérabilité (« des machines comprennent des machines »),
- faire sens pour l'utilisateur (« des humains comprennent des humains via des machines »).

C'est dans cette dernière approche que s'inscrit notre projet, celle d'un « Web cognitivement et socialement sémantique » (Cahier *et al.*, 2004).

La mise en œuvre du Web Sémantique nécessite, selon son inventeur Tim Berners-Lee, de disposer d'une « ontologie » du domaine considéré, c'est-à-dire d'un modèle consensuel des connaissances de ce domaine. Force est de constater que dans la plupart des domaines, on dispose de plusieurs « ontologies » souvent en contradiction. Cela s'explique par le fait que dans les professions d'expertise (ingénieur, chercheur, médecin, avocat...) la connaissance se construit de manière non pas consensuelle mais polémique. Si l'expert est considéré comme tel, c'est parce qu'il a pris une certaine distance vis-à-vis du savoir de référence qui lui a été enseigné.

Dans un « Web cognitivement et socialement sémantique », la pluralité des points de vue au sein d'une communauté n'est donc pas un frein à l'expression de la sémantique, mais au contraire le révélateur d'un travail de construction du sens.

Sylvie Calabretto (informatique)

L'idée d'organiser un corpus documentaire selon différents points de vue en fonction de différents usages rejoint nos travaux en collaboration avec la Caisse Nationale d'Allocations Familiales (CNAF) sur la multistucturalité des textes réglementaires.

Les textes réglementaires constituent une véritable matière première pour une organisation de ce type. Ces textes ont la particularité de donner à la fois des règles, quasi-formelles, ainsi qu'une quantité de cas particuliers nécessitant une expertise approfondie.

Ainsi, des systèmes d'information permettent aujourd'hui de gérer de manière automatique 80% des cas (pour dix millions d'allocataires), alors que 20% sont gérés par des experts en s'appuyant sur les textes.

De ces deux usages du corpus (par la machine et par l'homme) découlent deux structures concurrentes appelées respectivement « structure sémantique » et « structure logique », sachant qu'aucune ne peut être obtenue automatiquement à partir de l'autre.

Une thèse en cours (celle de Nouredine Chatti) explore la possibilité de visualiser et de manipuler conjointement ces différentes structures. L'un des principaux enjeux est de gérer plus efficacement la cohérence entre les deux structures au fur et à mesure des évolutions de la législation.

Bruno Helly (archéologie, histoire, géographie)

Il apparaît essentiel de croiser des points de vue complémentaires sur le matériel archéologique en particulier dans le cas des stèles et monuments funéraires. Ceux-ci constituent la partie visible, on dit en grec les *sēmata* des sépultures, qui comprennent aussi la tombe elle-même. *Sēma* (au pluriel *sēmata*) désigne le signe, la marque, de quelque chose.

Une stèle comme monument archéologique est en première analyse à la fois un produit matériel relevant de techniques de fabrications mises en oeuvre par des artisans, les tailleurs de pierre, les marbriers, et un "support de significations" (*Bedeutungsträger*), que l'on peut relier à des contextes d'intentions très différents les uns des autres.

Les différents composants des monuments funéraires sont normalement traités séparément les uns des autres :

- provenance et taille de la pierre,
- sculpture,
- peinture,
- architecture,
- inscriptions (en particulier, étude des noms propres)...

Chaque type de composant, dans la littérature archéologique, a fait le plus souvent l'objet de travaux séparés, insérés dans des bibliographies qui ne se recoupent pas. Il s'agit désormais de les étudier dans leurs relations et comme une totalité.

On n'apprend évidemment moins sur ces nécropoles que par des fouilles archéologiques, mais la perspective choisie permet d'apprendre beaucoup sur les monuments eux-mêmes et leur utilisation par les hommes.

3.3. Diachronie

3.3.1. Etude de cas

Notre troisième et dernière étude de cas porte sur la Chronique des fouilles et découvertes archéologiques (du Bulletin de Correspondance Hellénique, revue de l'École française d'Athènes). Le rôle de cette chronique est de signaler chaque année les « nouveautés » archéologiques dans le monde grec.

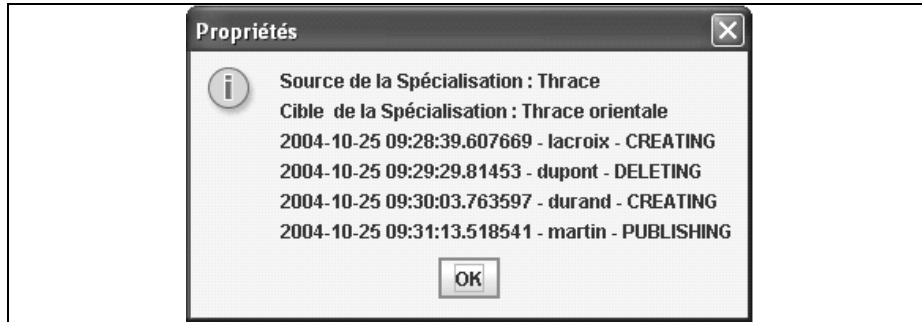


Figure 8. Historique des remaniements d'un arc du graphe (capture d'écran du système Porphyry)

La Chronique est diffusée sous forme de livraisons annuelles (depuis 1920) constituées d'éléments documentaires intelligibles indépendamment les uns des autres, structurés suivant une hiérarchie de sections donnant une impression de « zoom » géographique.

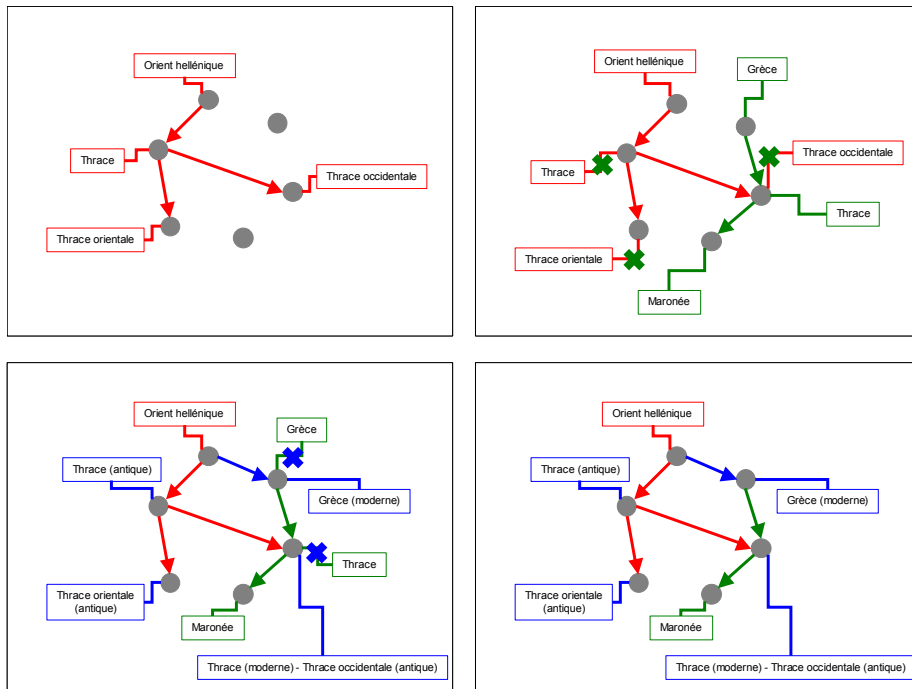


Figure 9. Visualisation de l'historique du réseau (maquette)

Cette structure bien que commune à l'ensemble des livraisons, a été remaniée un certain nombre de fois au cours de son histoire.

1^{er} ACTE : Le champ d'étude de la Chronique s'étend à l'Orient Hellénique, c'est-à-dire par rapport à la géographie antique.

2^{ème} ACTE : Le champ d'étude de la Chronique est ramené à la Grèce moderne. La « Thrace orientale », aujourd'hui en Turquie, n'en fait donc plus partie. La « Thrace » d'aujourd'hui est donc la « Thrace occidentale » d'autrefois. De plus, on précise la structure en indiquant que la « Thrace » contient « Maronée ».

3^{ème} ACTE (FICTIF) : Pour réaliser une bibliothèque numérique comprenant l'ensemble du corpus, on renomme chacun de ces espaces en précisant s'il s'agit du nom antique ou du nom moderne.

EPILOGUE : La structure est dans son état actuel... Jusqu'au prochain changement.

3.3.2. *Discussion interdisciplinaire*

Aurélien Bénéel (informatique)

Trop souvent, la science nous est présentée comme une « boîte noire », alors que comme nous le fait remarquer Bruno Latour (Latour, 1987), elle peut devenir « boîte de Pandore » si l'on s'intéresse à la « Science en action », c'est-à-dire à la science en train de se faire.

Une lecture philosophique de la Science, nous apprend en effet qu'une « loi scientifique » ne peut être vraie que par hypothèse, qu'elle est une construction de l'esprit du scientifique pour expliquer le monde (cf. concepts *a priori* chez Kant), et que par là, elle peut être réfutée lorsque ses prédictions sont mises en défaut par l'expérience (Popper, 1935).

Une lecture plus sociologique, nous montre que l'acceptation de ces « lois » repose sur un certain nombre de paradigmes : du « non-dit », des règles et des croyances absentes du discours scientifique. Le remplacement d'une loi scientifique par une autre cacherait le plus souvent un glissement d'une métaphysique vers une autre (Kuhn, 1962).

Dans chacune de ces lectures, la science apparaît comme construction dynamique de modèles explicatifs du monde. Par conséquent, nous défendons l'idée que des systèmes informatiques destinés aux chercheurs doivent leur permettre de réaliser cette modélisation dynamique. Ceci est sans doute à rapprocher de la position de Bruno Bachimont visant à assigner aux techniques de l'intelligence artificielle non plus une vocation de « représentation » au sens fort, mais celle d'une « écriture dynamique à interpréter » (Bachimont, 1999).

Andrea Iacovella (historiographie, archéologie)

Pour l'historiographie, « l'histoire a une histoire » et donc autant de sens actualisés. En effet, selon Michel de Certeau, l'histoire naît de la confrontation entre un passé et un présent, c'est-à-dire entre ce qui organisait la vie ou la pensée et ce qui permet aujourd'hui de la penser. Par conséquent, il n'existe pas *un* sens historique mais une *série* de sens historiques (De Certeau, 1975).

Notons que, non seulement, il y a une succession d'interprétations, mais que cette succession elle-même fait sens. Ainsi, quand nous défendons une certaine traçabilité des discours scientifiques, celle-ci n'est pas destinée uniquement à devenir un objet d'étude pour l'historien des sciences, mais également à constituer un vecteur de compréhension pour le chercheur de la discipline, l'apprenant, et surtout pour le public. En effet, il s'agit d'inventer un nouveau type de « vulgarisation » qui ôterait à ce mot son caractère péjoratif. Autrement dit, une telle approche participe de la mise en place d'une meilleure interaction entre le débat scientifique et le débat civique.

3. Bilan et perspectives

Au fil de cet article, nous avons vu comment un « espace de manipulation de traces » permettait trois modes de construction du sens : l'intertextualité, l'intersubjectivité et la diachronie. Pour mettre à l'épreuve cet arsenal conceptuel, nous nous sommes fait l'écho des débats tenus depuis plusieurs années dans le cadre du réseau transdisciplinaire entre informatique et sciences historiques. Nous avons pu à cette occasion montrer comment ce modèle conceptuel était largement corroboré par les réflexions épistémologiques des experts du réseau.

L'étape suivante consiste à confronter les prototypes issus de ces modèles à la pratique quotidienne des experts. Dans ce but, nous avons mis en place des « plateformes d'expertise » autour de communautés d'experts, de leurs corpus documentaires et de leur problématique. Or, ce type d'expérimentation « en grandeur nature » et sur le moyen terme n'est pas sans poser problème. Le premier type de difficultés, que nous ne ferons qu'évoquer ici, concerne l'organisation humaine du dispositif (Comment, entre autres, assurer l'installation, la formation, le support technique ?). Le second type de difficultés concerne le « bruitage » des résultats d'expérience par des imperfections de l'outil de nature purement ergonomique. Sans pour autant remettre en cause la nécessité des plateformes d'expertise, nous envisageons depuis peu de mettre en place des dispositifs d'expérimentation plus ponctuels dans le temps et l'espace et portant exclusivement sur l'ergonomie des interfaces de visualisation et de manipulation.

Notons que ces problèmes d'ergonomie ne sont pas l'exclusivité des prototypes de recherche. Quelque soit le système utilisé, le fait que l'on soit aujourd'hui réduit à imprimer un document numérique dès que l'on souhaite se « l'approprier », nous

semble révélateur de l'urgence de ce type d'études. En particulier, nous envisageons d'explorer les métaphores du *feuilletage* et de *l'annotation* qui, bien que très présentes en informatique, nous semblent utilisées bien en deçà de leurs capacités évocatoires.

Les travaux de recherche présentés dans cet article sont financés par le *Programme Société de l'Information 2001-2004* (CNRS STIC-SHS). Ils sont portés par le réseau transdisciplinaire ARTCADHi-CNRS⁸.

4. Bibliographie

4.1. Positionnement

Bachimont B., « L'intelligence artificielle comme écriture dynamique : de la raison graphique à la raison computationnelle », In : Petitot J. (Ed.), *Au nom du sens*, Paris : Grasset, 1999. p.290-319. Disponible sur Internet (consulté le 22/10/04) : <http://www.utc.fr/costech/v1/docs/intelligence_artificielle.pdf>

Cahier J.P., Zacklad M., Monceaux A., « Une application du Web socio-sémantique à la définition d'un annuaire métier en ingénierie », In : *Actes du colloque Ingénierie des Connaissances, Lyon, Mai 2004*.

Charlet J., Daigne M., Leroux V., « Ingénierie des patrimoines informationnels de l'établissement de santé », In : *Document numérique, volume 3, n°3-4, décembre 1999*. p.81-99.

De Certeau M., *L'écriture de l'histoire*, Paris : Gallimard, 1975, 358 p.

De Certeau M., *La faiblesse de croire*, Paris : Seuil, 1987.

Kuhn T. S., *La structure des révolutions scientifiques*, Paris : Flammarion, 1983, 284 p. (Note : Édition originale en anglais publiée en 1962, revue et augmentée en 1970)

Latour B., *La science en action*, Paris : Gallimard, 1995, 663 p. (Note : Édition originale en anglais publiée en 1987).

Nelson T.H., « Xanalogical Structure Needed Now More Than Ever », In : *ACM Computing Surveys, Volume 31, Issue 4, 1999, ACM Press, 1999*, Article 33. Disponible sur Internet (consulté le 22/10/04) : <http://www.cs.brown.edu/memex/ACM_HypertextTestbed/papers/60.html>

⁸ ARTCADHi : Atelier de Recherches Transdisciplinaires sur la Construction du sens en Archéologie et autres Disciplines Historiques. <<http://www.porphyry.org>>

Popper K. R., *La logique de la découverte scientifique*, Paris : Payot, 1973, 480 p. (Note : Édition originale en allemand publiée en 1935, puis nouvelle édition en anglais publiée en 1959 et complétée en 1960, 1962, 1965 et 1968)

Ricœur P., *Du texte à l'action : Essais d'herméneutique II*, Paris : Seuil, 1986, 452 p.

4.2. *Projet Porphyry*⁹

2004

Bénel A., « Expression du point de vue des lecteurs dans les bibliothèques numériques spécialisées », In : *Actes du Colloque International sur le Document Numérique, "Approches sémantiques sur le document numérique"*, La Rochelle, 22-25 juin 2004. p.???

Accary T., Bénel A., Calabretto S., Iacovella A., « Confrontation de points de vue sur des corpus documentaires : Le cas de la modélisation du temps archéologique », In : *Actes du 14ème Congrès Francophone AFRIF-AFLA de Reconnaissance des Formes et Intelligence Artificielle*, Toulouse, 28-30 janvier 2004. p.197-205.

Bénel A., Calabretto S., « Ontologies... déontologie : réflexion sur le statut des modèles informatiques », In : *Rivista "Linguistica Computazionale"*, Volume 20-21. Pisa/Roma : Istituti Editoriali e Poligrafici Internazionali, 2004. p.31-47.

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.), *Revue-texto.net : site de sémantique des textes*, mars 2004. 190 p. + Annexes. (Note : Thèse de doctorat en informatique soutenue à l'INSA de Lyon le 12 décembre 2003)

Abascal R., Beigbeder M., Bénel A., Calabretto S., Chabbat B., Champin P.-A., Chatti N., Jouve D., Prié Y., Rumpler B., Thivant E., « Documents à structures multiples », In : *Conference on Sciences of Electronic, Technologies of Information and Telecommunications*, Sousse, 15-20 mars 2004.

2003

Accary T., Bénel A., Calabretto S., « Modélisation de connaissances temporelles en Archéologie », In : *Actes des Journées francophones d'Extraction et de Gestion des Connaissances, Revue des Sciences et Technologies de l'Information [RSTI]*, Volume 17, Numéro spécial. Paris : Lavoisier - Hermès Sciences, 2003. p.503-508.

Abascal R., Beigbeder M., Bénel A., Calabretto S., Chabbat B., Champin P.-A., Chatti N., Jouve D., Prié Y., Rumpler B., Thivant E., « Modéliser la structuration multiple des documents », In : *Actes de la Conférence Hypertexte et Hypermédia, "Créer du sens à l'ère du numérique"*, Paris, 24-26 septembre 2003. Paris : Hermès, 2003. p.253-258.

⁹ La plupart des articles sont disponibles « en ligne » à l'adresse : <http://www.porphyry.org/artcadhi/publications>

20

2002

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.

Bénel A., Calabretto S., Iacovella A., Pinon J.-M., « Porphyry 2001: Semantics for scholarly publications retrieval », In : *Proceedings of the thirteenth International Symposium on Methodologies for Intelligent Systems*, Lyon, June 26-29, 2002. Lecture Notes in Artificial Intelligence #2366. Berlin : Springer-Verlag. p.351-361.

2001

Bénel A., Egyed-Zsigmond E., Prié Y. Calabretto S., Mille A., « Truth in the Digital Library: From Ontological to Hermeneutical Systems », In : *Proceedings of the fifth European Conference on Research and Advanced Technology for Digital Libraries*, Darmstadt, September 4-9, 2001. Lecture Notes in Computer Science #2163. Berlin : Springer-Verlag. p.366-377.

2000

Bénel A., Calabretto S., Pinon J.-M., Iacovella A., « Consultation de documents et sémantique : Application à des publications savantes », In : *Actes du second Colloque International Francophone sur l'Écrit et le Document*, Lyon, 3-5 Juillet 2000, Lausanne : PPUR. p.271-280.

Bénel A., Calabretto S., « Exploration de corpus de documents archéologiques à l'aide de théories algébriques », In : *Escuela interlatina de altos estudios en linguística aplicada, Segundo seminario, "Matemáticas y tratamiento de corpus"*, San Millan de la Cogolla, La Rioja, España, 19-23 Septiembre 2000, Logroño : Fundación San Millan de la Cogolla, 2002. p.343-350.

Bénel A., Calabretto S., Pinon J.-M., Iacovella A., « Vers un outil documentaire unifié pour les chercheurs en archéologie », In : *Actes du XVIIIe congrès INFORSID*, Lyon, 16-19 Mai 2000, Éditions INFORSID. p.133-145.

1999

Bénel A., Calabretto S., Pinon J.-M., « Indexation "sémantique" de documents archéologiques », In : *Deuxième colloque du chapitre français de l'ISKO, "L'indexation à l'heure d'Internet"*, Lyon, 21-22 Octobre 1999, Éditions ISKO-France, Imprimeur ENSSIB, 2001. p.145-152.