

---

# Lecture, interprétation et constitution de corpus en Sciences Humaines

## Aspects épistémologiques et computationnels<sup>1</sup>

**Aurélien Bénel**

*Laboratoire d'Ingénierie des Systèmes d'Information (INSA de Lyon)  
Ecole française d'Athènes  
Aurelien.Benel@insa-lyon.fr*

---

*RÉSUMÉ. Notre étude porte sur les dossiers numériques dans les Sciences Humaines et plus particulièrement en Archéologie. En nous rattachant à la tradition herméneutique (c'est-à-dire interprétative) de ces sciences, nous essayons de montrer pourquoi le dossier, dans sa lecture, son interprétation et sa constitution, matérialise l'outil idéal du travail de ces chercheurs. De ces méthodes, nous tentons d'extraire des « patrons de conception » pour des « systèmes d'assistance à l'interprétation » ; c'est-à-dire des systèmes informatiques qui permettraient d'instrumenter un travail humain de construction de sens.*

*ABSTRACT. This paper deals with the use of digital dossiers in Humanities and especially in Archaeology. From an hermeneutical point of view, the bundle of fragments in the “dossier”, by being read, interpreted and gathered, could be seen as the ultimate tool for these researchers. From these theoretical aspects, we try to extract several “design patterns” for “Interpretation Assistance Systems”, i.e. computer systems designed to assist human sense-making processes.*

*MOTS-CLÉS : Herméneutique, Epistémologie, Explication/Compréhension, Systèmes d'Assistance à l'Interprétation, Annotation, Lecture/Ecriture.*

*KEYWORDS: Hermeneutics, Epistemology, Explanation/Understanding, Interpretation Assistance Systems, Annotating, Reading/Writing.*

---

---

<sup>1</sup>Par « computationnel », nous entendons ce qui a trait au calcul mécanisé (automatique), c'est-à-dire à l'informatique.

## **1.Introduction**

Il apparaît de plus en plus clairement que l'instrumentation du travail des experts passe par le *dossier numérique*. Comme l'a analysé Jean Charlet pour le domaine 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).

En écho, nous proposons une étude de cas dans un autre domaine d'expertise : celui de la recherche en Sciences Humaines, et plus particulièrement de l'Archéologie. L'objectif est double. Il s'agit d'une part de comprendre pourquoi l'approche du dossier (on parlera ici de *corpus*) est compatible avec le travail de ces chercheurs, à l'opposé de l'approche factuelle par exemple. D'autre part, il s'agit de proposer quelques choix de conception (implantés dans notre prototype ou en voie d'implantation) qui pourraient être réutilisables, nous le pensons, dans d'autres systèmes de gestion de dossier numérique.

Nous essaierons dans un premier temps de replacer notre étude dans les besoins qui l'ont suscitée, ainsi que dans l'approche méthodologique que nous avons choisie. Ensuite, nous verrons en quoi, les méthodes elles-mêmes des Sciences Humaines recommandent la lecture, l'interprétation et la constitution de corpus documentaires par différents acteurs. Enfin, nous verrons comment de ces méthodes, il est possible de bâtir des « patrons de conception » pour des « systèmes d'assistance à l'interprétation » ; c'est-à-dire des systèmes informatiques qui, tout en restant les plus discrets possibles, permettent d'instrumenter un travail humain de construction de sens.

## **2.Etude des besoins**

### ***2.1.Contexte et motivation***

Le projet émane de l'Ecole française d'Athènes. Par sa maison d'édition et sa riche collection documentaire (bibliothèque, photothèque, "planothèque", archives manuscrites), elle occupe une place de choix dans la construction et la transmission des savoirs en archéologie. C'est dans ce contexte que nous tâchons de concevoir et de mettre en place des outils informatiques adaptés à la gestion de ces savoirs.

Plusieurs dizaines d'années après les grands projets de bases factuelles, de systèmes experts et autres « banques de données », le chercheur en archéologie a toujours, comme principaux outils, les publications savantes. Jean-Claude Gardin, instigateur de certains de ces projets, reconnaît lui-même le peu d'intérêt que rencontrèrent ses banques de données archéologiques et interprète ces échecs

comme étant dus à la difficulté de distinguer en archéologie les « faits », des conclusions ou des interprétations (Gardin, 1986). Pour aller plus loin, on pourrait se demander si le paradigme computationnel, qui a pourtant eu le succès qu'on lui connaît dans les entreprises, ne serait pas, tel quel, inapplicable à un certain nombre de domaines. La Médecine, le Droit, les Sciences de l'Homme et de la Société, sont autant de domaines où la recherche en informatique a dû s'adapter au seul support de la connaissance reconnu valide par ses praticiens : le document. Sans même requérir l'argument épistémologique de l'inadéquation de la formalisation aux Sciences de l'Homme, il est impossible, comme le concède Jean-Claude Gardin<sup>2</sup>, de parler de bases de données archéologiques sans s'interroger sur la reconnaissance de ces données par l'ensemble de la communauté des chercheurs en archéologie. En effet, à la différence des bases de données, le document daté, authentifié et soumis à un comité éditorial est en adéquation totale avec une science moderne de type constructiviste, c'est à dire où la scientificité repose sur la « réfutabilité » d'hypothèses par l'expérience (Popper, 1959) et/ou par les pairs (Kühn, 1962).

Si les documents semblent adaptés à la pratique scientifique, il en est autrement des systèmes documentaires courants. En effet, conçus pour des bibliothèques grand-public, la plupart reposent sur une indexation effectuée par des tiers selon une structure *a priori* (ontologie, thésaurus, liste d'autorité). Or, structurer le corpus de documents – et par là le domaine – à la place du chercheur, c'est nier son expertise. En effet, c'est la mission même du chercheur de trouver une structure à son objet d'étude et de tester la validité de cette structure. Par conséquent, un système documentaire adapté aux chercheurs ne devrait pas être basé sur des « méta-données » fixes, mais être le support d'une activité exploratoire de structuration.

## **2.2. Méthodologie**

Instrumenter un travail n'est jamais une opération neutre. Comme le rappelle Bruno Bachimont (Bachimont, 1999a) n'importe quel outil (du marteau au système informatique) détermine par sa structure des usages possibles (ce qui n'empêche pas des usages déviants). La question de l'adéquation de l'outil au travail à effectuer est donc déjà sensible au niveau d'une entreprise. Elle devient flagrante dans le cas de la recherche publique, pour laquelle, il serait tout à fait déplacé de la part d'une instance hiérarchique d'imposer l'utilisation d'un outil.

Pour proposer des outils adéquats à la recherche scientifique, deux voies se présentent. La première consiste en un travail « d'anthropologie », c'est à dire d'immersion dans une communauté de chercheurs et d'élaboration d'un modèle du travail à partir d'observations et d'entretiens (plus ou moins dirigés). La seconde repose sur l'étude des textes méthodologiques qui guident le travail de ces

---

<sup>2</sup> « Par quels mécanismes obtiendra-t-on que l'accord initial [...] engageant une population de chercheurs limitée dans l'espace et le temps, s'étende ensuite de façon quasi-statutaire [...] ? » (Gardin, 1984)

chercheurs. Dans les deux cas, notre étude visant à la modélisation informatique d'outils pour la construction de savoirs scientifiques ne peut s'abstraire d'une théorie de la connaissance, c'est à dire d'une épistémologie. Ce « détour » par la philosophie n'est pas le fruit d'une quelconque fantaisie mais la conséquence directe de notre domaine d'application.

Notre exploration de la première voie nous a permis de corroborer des études existantes mettant en évidence la lecture comme réécriture (O'Hara *et al.*, 1998) et l'importance des interactions sociales en ce qui concerne la découverte, l'interprétation et la diffusion des connaissances (Paepcke, 1996). Cependant, nous nous sommes vite retrouvés face à des problèmes méthodologiques majeurs que, n'étant pas anthropologues, nous n'avons pas réussi à résoudre. Aussi, nous avons préféré approfondir la deuxième voie consistant à modéliser dans un système informatique les méthodes données par les épistémologues.

### 3.Aspects épistémologiques

Afin de dresser un aperçu de l'épistémologie des Sciences Humaines, il nous est nécessaire de préciser tout d'abord ce que l'on entend par « Sciences Humaines ». L'œuvre de Wilhelm Dilthey (fin XIX<sup>e</sup> s. - début XX<sup>e</sup> s.) est en ce sens très éclairante (Ricoeur, 1986). Elle propose en effet un critère de démarcation entre *les Sciences de la Nature* et les Sciences Humaines (qu'il appelle *Sciences de l'Esprit*). Pour lui, ce critère repose sur la dialectique entre *explication* et *compréhension*. L'explication serait le mode de construction des Sciences de la Nature, tandis que la compréhension serait celle des Sciences de l'Esprit.



**Figure 1.** L'interprétation d'après Dilthey comme compréhension pure (disjointe de l'explication)

L'explication concerne le « comment », la compréhension le « pourquoi ». Alors que le premier est l'étude d'un *objet* indépendamment de l'observateur (objectivité), la seconde est l'étude d'un *sujet* (une personne) par un autre sujet (subjectivité). L'Explication nécessite une mise à distance (*distanciation*). A l'inverse la compréhension (« prendre avec ») requiert une appropriation (*actualisation*). La

première manipule de la *signification*, propriété de signes *décontextualisés*<sup>3</sup>, à la différence de la seconde qui manipule du *sens*, propriété de signes en *contexte* et en *situation* (cf. François Rastier<sup>4</sup> sur la dialectique Sens/Signification ; [Ducrot *et al.*, 1972] sur la dialectique Contexte/Situation).

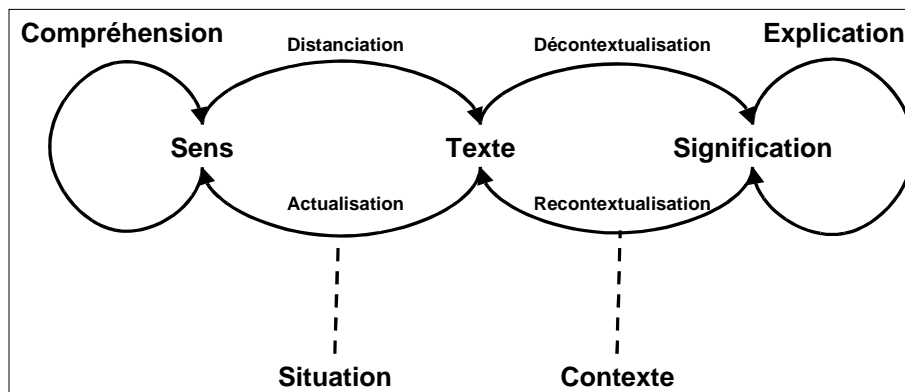
On entend par *contexte d'un signe*, étymologiquement parlant, ce qui est « tissé avec », c'est-à-dire l'ensemble des signes qui l'entourent. Un *texte* est ainsi, pour reprendre son étymologie, un « tissu ». La *situation*, quant à elle, étend la notion de *contexte* à un « tissu » également non linguistique (temps, lieu, personne...). Pour un texte donné, peuvent être considérées à la fois les situations d'écriture et de lecture. La première, par exemple, comprend la culture et l'histoire à la fois de l'auteur et de la société auxquels le texte appartenait.

Afin d'éviter de plonger les Sciences Humaines dans une subjectivité débridée, Dilthey choisit de placer la compréhension dans la tradition méthodologique de *l'Herméneutique*, c'est à dire la science de l'interprétation. En effet, cette science destinée au départ à l'étude des textes sacrés, avait posé la question des limites de l'interprétation et proposé un certain nombre de réponses : étudier la situation dans laquelle un texte a été écrit, comparer un texte avec des textes contemporains ou antérieurs, avec d'autres versions du même texte, avec des commentaires postérieurs, multiplier les interprétations et engager le débat, ou en dernier recours s'en remettre à la *tradition* d'interprétation, à *l'autorité*, ou au *préjugé* pour se forger une opinion provisoire.

---

<sup>3</sup> Par exemple, dans le cadre de référence de la Physique classique, « masse » et « vitesse » possèdent une signification fixe quels que soient leurs contextes d'apparition.

<sup>4</sup> « [...] la signification est une propriété assignée aux signes, le sens une propriété des textes. [...] la signification résulte en effet d'un processus de décontextualisation [...]. En revanche le sens suppose une contextualisation maximale aussi bien par la langue (le contexte, c'est tout le texte) que par la situation (qui se définit par une histoire et une culture, au-delà du *hic et nunc* de la pragmatique). » (Rastier, 1998) p.7-8



**Figure 2.** L'interprétation d'après Ricoeur comme compréhension et explication mises en relation par le texte.

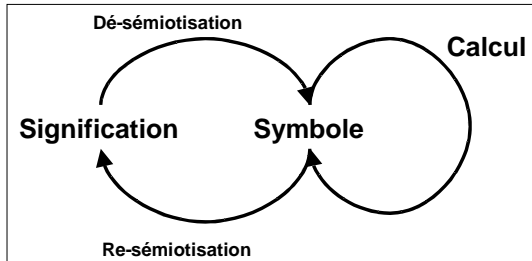
En résumé, avec Dilthey, apparaissait une épistémologie assez claire des Sciences de l'Esprit, épistémologie guidée par une herméneutique de la compréhension, totalement disjointe du processus explicatif (cf. **Figure 1**). Or, ce modèle va être remis en cause avec la naissance du structuralisme en linguistique au début du XX<sup>e</sup> siècle et surtout son extension à toutes les spécialités des Sciences Humaines dans les années 1960 et 1970. En effet, ce mouvement à tendance clairement explicative va s'avérer d'une fécondité remarquable (en anthropologie, psychanalyse, critique littéraire, étude de la grammaire, etc.) réfutant l'omission de la dimension explicative dans l'épistémologie des Sciences Humaines. Par contre, la tentation de réduire les Sciences Humaines à la seule explication s'est avérée à l'expérience comme désastreuse.

La thèse centrale de Paul Ricoeur dans son deuxième essai d'herméneutique (Ricoeur, 1986) est de rassembler sous la méthode herméneutique les dimensions de la compréhension et de l'explication ; de considérer l'interprétation non plus comme la seule compréhension mais comme la dialectique même de la compréhension et de l'explication. En effet, la distanciation nécessite l'appartenance préalable à des sujets, et à l'inverse il n'existe pas de compréhension de soi sans médiation par des signes, symboles ou textes.<sup>5</sup> En résumé, pour Ricoeur, le *texte*<sup>6</sup> est le trait d'union entre les

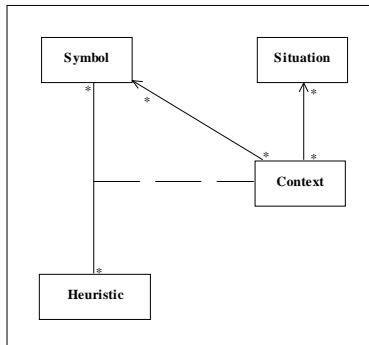
<sup>5</sup> « Ce combat sur deux fronts contre une réduction de la compréhension à l'intropathie et une réduction de l'explication à une combinatoire abstraite m'amène à définir l'interprétation par cette dialectique même de la compréhension et de l'explication au niveau du "sens" immanent au texte. » (Ricoeur, 1986) p.37-38

<sup>6</sup> « [...] quelle peut être la tâche première de l'herméneutique ? Elle est selon moi, de chercher dans le texte lui-même, d'une part la dynamique interne qui préside à la structuration de l'œuvre, d'autre part la puissance de l'œuvre de se projeter hors d'elle-même et d'engendrer un monde qui serait véritablement la "chose" du texte. Dynamique interne et projection externe constituent ce que j'appelle le travail du

processus de compréhension et d'explication ; l'ensemble de ces deux processus étant régi par les méthodes de l'herméneutique (cf. **Figure 2**).



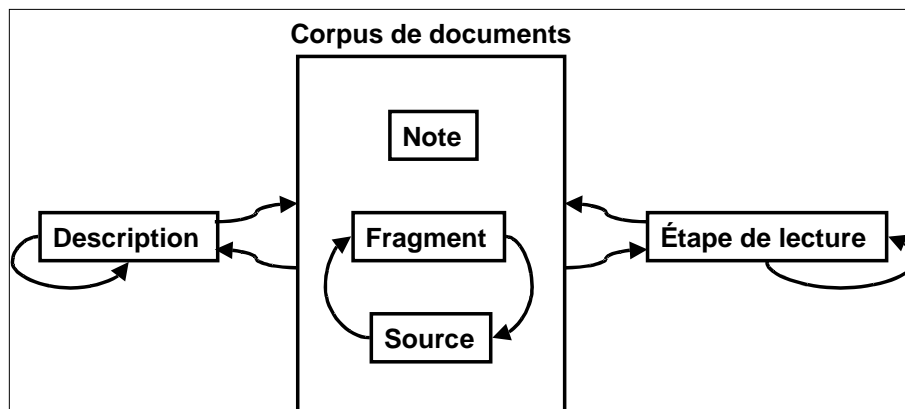
**Figure 3.** *L'explication par le calcul d'après Kanellos : le passage obligé de la signification au symbole (et inversement)*



**Figure 4.** *La « trace » comme union d'un symbole, de son contexte et d'indices de situation*

---

texte. C'est la tâche de l'herméneutique de reconstruire ce double travail du texte. » (Ricœur, 1986) p.36.



**Figure 5.** *Système Porphyre : Liens contextuels entre traces*

Reste à définir la place du calcul (c'est-à-dire de l'informatique) dans notre modèle. Le calcul apparaît clairement comme faisant part du processus d'explication. Cependant, des précisions s'imposent. En effet, alors que l'explication, comme nous l'avons vu, porte sur des significations, le calcul porte sur des *symboles* (au sens mathématique) c'est-à-dire des objets de calcul sans aucune signification. Ioannis Kanellos parle de « forme décorrélée de tout contenu » (Kanellos *et al.*, 2000); Bruno Bachimont de signe *autothétique* (étymologiquement, « qui porte soi-même »), c'est-à-dire « qui ne représente rien sinon sa propre effectivité » (Bachimont, 1999b). Là où les choses se compliquent, c'est que les usagers des systèmes informatiques attribuent une signification aux symboles (et même plus puisqu'ils sont, selon l'expression de François Rastier<sup>7</sup>, « condamnés au sens »). Bruno Bachimont considère que le symbole autothétique se voit surdéterminé jusqu'à devenir *orthothétique* (étymologiquement, « qui porte exactement »), c'est-à-dire un enregistrement de la parole telle que l'écriture alphabétique<sup>8</sup>. Ioannis Kanellos en vient à définir deux processus humains à savoir la *dé-sémiotisation* et la *re-sémiotisation* qui permettent de passer respectivement de la signification au symbole et du symbole à la signification (Kanellos *et al.*, 2000).

<sup>7</sup> « [...] les phrases réputées absurdes, voire asémantiques peuvent toujours être interprétées [...] on peut lire *Finnegan's Wake* même dans les passages où aucun des mots ne figurent au dictionnaire [...] C'est là une allégorie du péché originel, ou du moins de la condition humaine : nous sommes condamnés au sens. » (Rastier, 1991) p.212-213

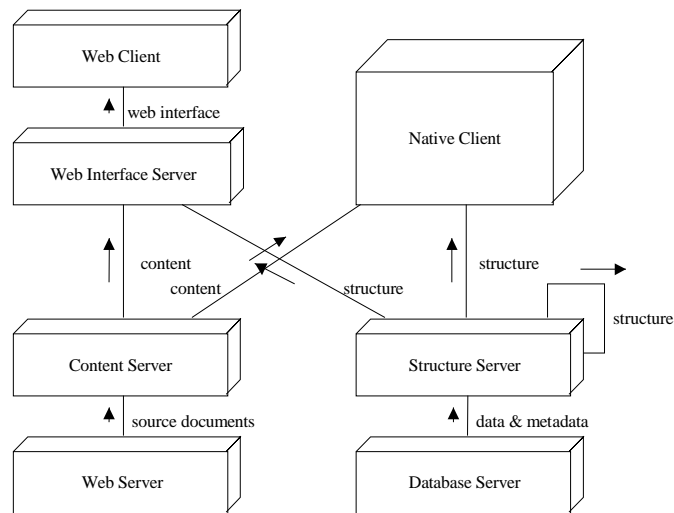
<sup>8</sup> D'où son projet de considérer l'intelligence artificielle non plus comme une modélisation au sens fort mais comme une écriture dynamique à interpréter (Bachimont, 1999a). Notons qu'il s'agit de l'approche inverse de celle de Jean-Claude Gardin qui, pour éviter les écueils supposés de l'interprétation, préconisait d'appliquer le formalisme des systèmes experts aux publications savantes en Sciences Humaines (Gardin, 1986).

Prenons un exemple. Considérons un chercheur travaillant sur le dieu Hermès. En cherchant dans un corpus de textes le nombre d'occurrences de l'expression régulière « HERM\* », il dé-sémiotise une signification en une suite de symboles informatiques. Le système calcule une liste de réponse. Le chercheur effectuera ensuite une re-sémiotisation, en interprétant par exemple la faible fréquence des occurrences comme une certaine discrétion du dieu dans le panthéon. Notons que dé-sémiotisation et re-sémiotisation ne sont jamais neutres. Par exemple le fait de compter toutes les occurrences de « HERM\* » relève du choix du chercheur de considérer l'emploi métonymique de « hermès » (buste quelconque surmontant un support quadrangulaire) comme également représentatif de l'influence du dieu.

Ainsi, le processus explicatif se décompose en un processus humain de dé-sémiotisation permettant de transformer la signification en un symbole sémiotiquement neutre, puis d'un processus automatique de calcul, et enfin d'un processus humain de re-sémiotisation afin de donner une signification au résultat du calcul (cf. **Figure 3**).

Pour résumer, les processus humains de construction de sens que nous avons étudiés nécessitent des symboles, ces symboles, tout d'abord re-sémiotisés, sont ensuite mis en contexte, et enfin ces symboles contextualisés sont mis en situation.

#### 4.Aspects computationnels



**Figure 6.** *Système Porphyre : Architecture logicielle*

Des considérations épistémologiques précédentes, nous pensons qu'il est possible d'en extraire une sorte de canevas ou de patron (Gamma *et al.*, 1994) permettant de guider la conception des outils informatiques pour les Sciences Humaines.

#### 4.1. Traces

Le sens d'un document n'est pas dans sa structure (quelle qu'elle soit) mais dans les conditions de son utilisation et donc dans les conditions de sa lecture. Il est impossible de trouver *le* sens d'un document mais plutôt *des* sens. Toute lecture est donc une réécriture. Un système adéquat devrait permettre de garder trace des parcours d'interprétation des lecteurs. Vannevar Bush, visionnaire de ce qui allait devenir l'hypertexte, décrivait d'ailleurs en 1945 un modèle basé sur des *traces* (ou pistes) de pensée (Bush, 1945). Cette métaphore de la « trace » est particulièrement riche et mérite que l'on s'y attarde un peu.

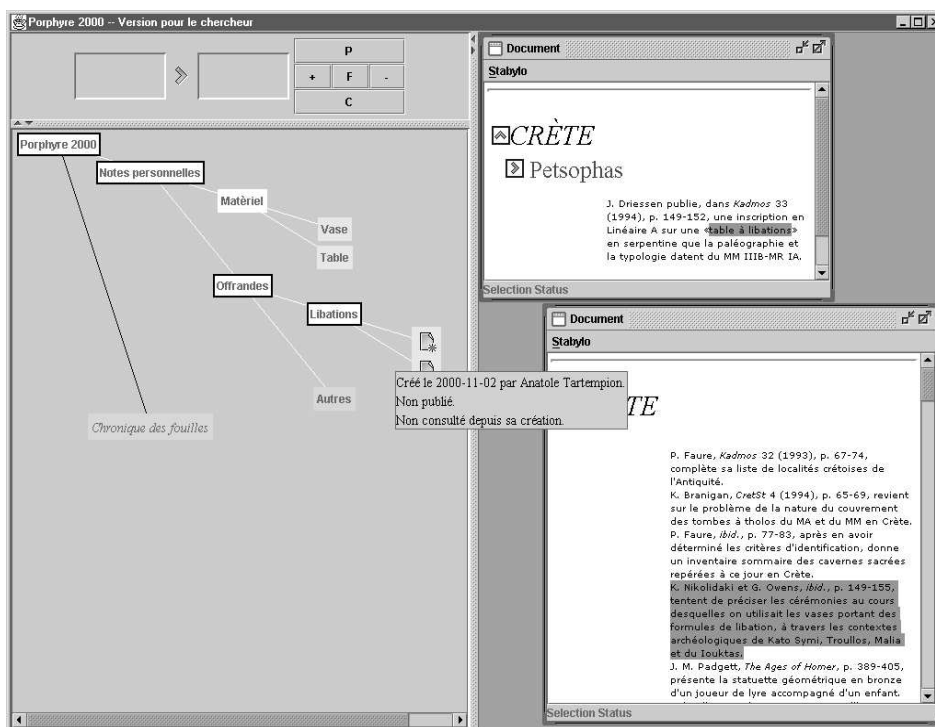


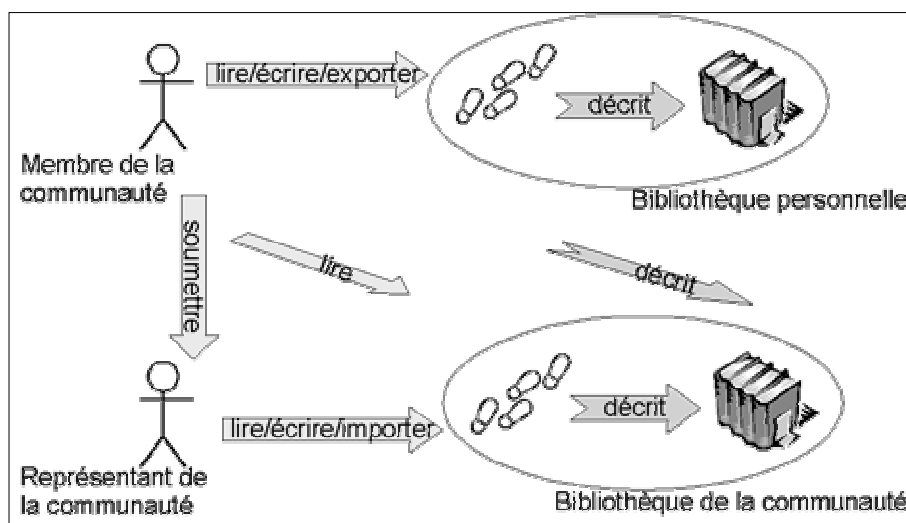
Figure 7. Système Porphyre : Copie d'écran du client natif

Tout d'abord, une trace, c'est la *forme* qui demeure quand le *contenu* a disparu. En effet, comme nous l'avons vu, l'ordinateur ne traite que de *symboles* dénués de signification. C'est à l'utilisateur qu'incombe la difficile tâche de donner un contenu

à ce qui n'est que formel. Ensuite, une trace est rarement seule, elle fait partie d'une piste. On retrouve ici la notion de *contexte*, c'est-à-dire d'autres traces reliées à elle. Enfin, on peut « suivre quelqu'un à la trace » mais aussi « suivre les traces de quelqu'un ». Dit autrement, il existe deux approches complémentaires : interpréter les traces d'un autre selon notre point de vue, ou tenter par la pensée de se mettre « à sa place ». On retrouve ici la notion de *situation*.

Si les symboles en contexte peuvent être « rentrés » dans une machine, il n'en est pas de même pour les situations. Il est impensable de faire rentrer dans un ordinateur des personnes, des lieux, des cultures, bref le monde. Aussi nous considérerons plutôt des « indices de situation », c'est-à-dire des informations<sup>9</sup> sur les différentes étapes de la « vie » d'un symbole contextualisé (date et auteur de la création d'un contexte, etc.).

Ainsi, notre approche théorique nous permet de proposer le patron de la « Trace » (cf. **Figure 4**). Pour un symbole donné, moyennant une heuristique (ex : locale, globale...), il est possible d'obtenir un contexte pointant vers un certain nombre d'autres symboles. Ce contexte est associé à des indices de situation. Notons qu'une trace peut être le résultat de l'agrégation d'autres traces, l'agrégation étant une mise en contexte particulière.



**Figure 8.** La "publication" comme processus d'intégration des traces lectoriales

<sup>9</sup> Strictement parlant, ces indices sont eux mêmes des symboles qui nécessiteraient une mise en contexte et en situation. Cependant, pour éviter une régression à l'infini, on supposera qu'ils seront tels-quels convenablement interprétés.

Dans notre prototype<sup>10</sup> (appelé *Porphyre*) ces traces (cf. **Figure 5**) peuvent être des documents : *notes*, *sources* ou *fragments*. Elles peuvent également être des *corpus* de documents, des *descriptions* semi-formelles de ces corpus (à l'aide de graphes acycliques de descripteurs), ou encore des *étapes de lecture*. Les descriptions sont mises en contexte entre-elles suivant les relations de généralisation/spécialisation qui existent entre leurs graphes respectifs. Les étapes de lecture, pour leur part, sont mises en contexte par les relations de séquence indiquées dans des *parcours de lecture*.

Pour la gestion des documents et des corpus (en particulier pour les fragments), nous utilisons un *serveur de contenu* (cf. **Figure 6**). Celui-ci est à distinguer des *serveurs web* stockant les documents sources. Il permet de transformer les documents sources sous la forme voulue (extraction d'une partie d'image ou de texte, « surlignage » de tous les fragments associés à un document, etc.).

Pour la gestion des traces semi-formelles, nous faisons appel à un *serveur de structure* (cf. **Figure 6**). Celui-ci stocke ses graphes dans une base de données et offre la possibilité d'interroger des serveurs de structure distants.

Les traces peuvent être consultées (cf. **Figure 6**) soit par un *client web* par le biais d'un *serveur d'interface web* soit par le *client natif* (cf. **Figure 7**). Ce dernier destiné au lecteur expert lui permet d'effectuer une lecture-écriture des traces suivant la métaphore de *l'annotation*. Nous nous plaçons ainsi dans la prolifique descendance des travaux de Jacques Virbel et Bernard Stiegler autour des « Postes de Lecture Assistée par Ordinateur » (Stiegler, 2000).

#### 4.2. Publication

Il ne suffit pas de stocker des traces, encore faut-il les partager... La métaphore que nous choisisons est celle de la *publication* (cf. **Figure 8**). Cette « mise en public » est basée sur une relation de confiance concrétisée par une communauté. L'adhésion d'un expert à une communauté nécessite l'accord commun de l'expert et du représentant de la communauté. Membre de la communauté, l'expert peut lire les documents et les traces publiés par les autres membres. Dans une bibliothèque privée, il peut ajouter ses propres traces d'interprétation relatives à ses propres documents ou aux documents de la communauté.

Si l'expert le souhaite, il peut soumettre au représentant de la communauté une exportation totale ou partielle de ses traces et documents personnels pour les publier. Le responsable de la communauté peut alors choisir de les publier ou non, éventuellement en y apportant des modifications consenties par l'auteur.

---

<sup>10</sup>Nous ne reviendrons pas ici sur la description détaillée des fonctions du système. Celle-ci pourra être trouvée dans nos précédents travaux (Bénel *et al.*, 2000a ; Bénel *et al.*, 2000b ; Bénel *et al.*, 2001).

La figure du « représentant de la communauté » permet de gagner en généralité. Il incombe ensuite à chaque communauté de décider quelle organisation elle mettra derrière : anarchie, consensus, majorité, comité de lecture, ou dictature !

D'un point de vue informatique l'adhésion d'un membre (ou d'une communauté) à une autre communauté se traduit par l'inclusion virtuelle des traces du second dans celles du premier. Etant donné qu'un *serveur de structure* correspond (*grosso modo*) à une communauté, les requêtes adressées au *serveur* d'un membre (cf. **Figure 6**) sont suivant le cas traitées sur place et/ou déléguées aux serveurs des communautés auxquels il est abonné (et ainsi de suite).

Dans un souci d'ouverture du système, la *soumission* d'un ensemble de traces se fait en dehors du système (par l'intermédiaire d'un courriel par exemple). Les traces sont exportées par leur auteur dans un fichier XML<sup>11</sup>, le fichier est transmis au représentant de la communauté qui les importe ensuite dans une zone *brouillon* où il peut les modifier. Le représentant peut enfin publier (au sens propre) les traces, c'est-à-dire les transférer dans la zone *publique* de la communauté.

## 5. Conclusion

Au début de cet article nous nous interrogeons sur ce que pourrait être un outil informatique pour la recherche en Sciences Humaines. Notre étude épistémologique nous a permis de caractériser l'activité de ces sciences par la production de sens. Si le sens reste intangible, sa production par contre « laisse des *traces* » sur un *support*. L'outil informatique pourrait ainsi se voir attribuer la fonction d'être ce support de lecture et d'écriture de traces (ces traces étant des symboles en contexte et accompagnés d'indices de situation).

De cette notion de trace qui pourrait s'apparenter à un « patron de conception » pour les « systèmes d'aide à l'interprétation », nous avons donné différents exemples tirés de notre prototype : fragments de documents, descripteurs, étapes de lecture. Par ailleurs, nous avons montré comment ces traces et les structures qui les mettent en contexte pouvaient évoluer au sein d'une communauté. En reprenant la métaphore de la *publication* comme processus d'intégration itérative des traces des lecteurs, nous avons proposé un moyen d'authentifier une trace (date, auteur) et de la placer dans le cadre des règles d'une communauté donnée.

Ainsi, l'étude épistémologique des Sciences Humaines, nous a permis de définir les deux notions de *trace* et de *publication*. Par la structuration d'un corpus (évolutive et collaborative) qu'elles supposent, ces notions semblent être de bonnes illustrations du concept de « dossier numérique ». Nous pensons quant à nous, qu'au-delà d'une simple illustration, il s'agit d'un exemple prototypique. Autrement dit,

---

<sup>11</sup> « eXtensible Markup Language » : Format standard de représentation de données semi-structurées (souvent utilisé comme format d'échange entre systèmes d'information différents).

nous faisons l'hypothèse que nos notions (voire même certaines de leurs implémentations dans notre prototype) pourraient être utilisées dans d'autres cas de « dossiers numériques ». Dans cette perspective, nous espérons pouvoir expérimenter nos concepts et/ou notre système dans d'autres domaines d'expertise que celui des Sciences Humaines afin de préciser leur degré de généralité.

## 6. Bibliographie

- (Bachimont, 1999a) Bachimont B., « L'intelligence artificielle comme écriture dynamique : de la raison graphique à la raison computationnelle », Petitot J. (Ed.), *Au nom du sens*, Paris, Grasset, 1999.
- (Bachimont, 1999b) Bachimont B., « De l'hypertexte à l'hypotexte : les parcours de la mémoire documentaire », *Technologie, Idéologies, Pratiques (TIP)*, n° spécial « Mémoires collectives », 1999.
- (Bénel *et al.*, 2000a) Bénel A., Calabretto S., Pinon J.-M., Iacovella A., « Vers un outil documentaire unifié pour les chercheurs en archéologie », *Actes du XVIIIe Congrès sur l'Informatique des Organisations INFORSID'2000, Lyon, 16-19 mai 2000*, Toulouse, INFORSID Editions, p.133-145.
- (Bénel *et al.*, 2000b) Bénel A., Calabretto S., Pinon J.-M., Iacovella A., « Consultation de documents et sémantique : Application à des publications savantes », *Actes du IIe Colloque International Francophone sur l'Écrit et le Document CIFED'2000, Lyon, 3-5 juillet 2000*, Lausanne, PPUR Editions, p.271-280.
- (Bénel *et al.*, 2001) Bénel A., Egyed-Zsigmond E., Prié Y., Calabretto S., Mille A., Iacovella A., Pinon J.-M., « Truth in the Digital Library: From Ontological to Hermeneutical Systems », *Proceedings of the fifth European Conference on Research and Advanced Technology for Digital Libraries ECDL'2001, Darmstadt, 4-9 septembre 2001, LNCS #2163*. Berlin, Springer-Verlag Editions, p.366-377.
- (Bush, 1945) Bush V., « As we may think », *The Atlantic Monthly*, July 1945.
- (Charlet *et al.*, 1999) Charlet J., Daigne M., Leroux V., « Ingénierie des patrimoines informationnels de l'établissement de santé », *Document numérique*, volume 3, n°3-4, décembre 1999. p.81-99.
- (Ducrot *et al.*, 1972) Ducrot O., Schaeffer J.-M., *Nouveau dictionnaire encyclopédique des sciences du langage, s.v. « Situation de discours »*, Paris, Seuil, 1995 [1972], p.764-775.
- (Gamma *et al.*, 1994) Gamma E., Helm R., Johnson R., Vlissides J., *Design Patterns: Elements of Reusable Object-Oriented Software*, Boston, Addison Wesley, 1994.
- (Gardin, 1984) Gardin J.-C., « Les bases de données dans les sciences de l'antiquité : l'ajustement nécessaire des fins aux moyens », *Banques de données et sciences de l'antiquité*, 1984.
- (Gardin, 1986) Gardin J.-C., « Systèmes experts et publications savantes », *The Fifth British Library Annual Research Lecture*, 1986.

- (Kanellos *et al.*, 2000) Kanellos I., Thlivitis T., Léger A., « Indexation et recherche d'information imagée par le contenu : catégories de textes et catégories d'images dans le projet SemIndex », *Actes du troisième Colloque International sur le Document Electronique CIDE'2000, Lyon, 4-6 juillet 2000*, Paris, Europa Productions, p.171-183.
- (Kühn, 1962) Kühn T. S., *La structure des révolutions scientifiques*, Traduction française, Paris, Flammarion, 1983 [1962].
- (O'Hara *et al.*, 1998) O'Hara K., Smith F., Newman W., Sellen A., « Student readers' use of library documents: implications for library technologies », *Conference Proceedings on Human Factors in Computing Systems CHI'98*, New York, ACM Press, 1998
- (Paepcke, 1996) Paepcke A., « Digital libraries: Searching is not Enough. What we learned on-site », *D-Lib Magazine*, May 1996.
- (Popper, 1959) Popper K. R., *La logique de la découverte scientifique*, Traduction française, Paris, Payot, 1973 [1959].
- (Rastier, 1991) Rastier F., *Sémantique et recherches cognitives*, Paris, PUF, 2001 [1991].
- (Rastier, 1998) Rastier F., « Sens et signification », *Protée*, printemps 1998. p.7-18.
- (Ricœur, 1986) Ricœur P., *Du texte à l'action : Essais d'herméneutique II*, Paris, Seuil, 1986.
- (Stiegler, 2000) Stiegler B., « Annotation, navigation, édition électronique : Vers une géographie de la connaissance », *Ec/arts*, n°2, 2000.