



**HAL**  
open science

# L'ANALYSE FORMELLE DES EGODOCUMENTS DANS UN SYSTÈME INFORMATIQUE DE PRODUCTION DE RESSOURCES ÉLECTRONIQUES

Dominique Taurisson-Mouret

► **To cite this version:**

Dominique Taurisson-Mouret. L'ANALYSE FORMELLE DES EGODOCUMENTS DANS UN SYSTÈME INFORMATIQUE DE PRODUCTION DE RESSOURCES ÉLECTRONIQUES. Corpus en lettres et sciences sociales : des documents numériques à l'interprétation Colloque international et école d'été, Albi, 10-14 juillet 2006, organisé dans le cadre des Colloques d'Albi Langages et signification (CALs), 2006, Albi, France. pp.43-50. halshs-00096435

**HAL Id: halshs-00096435**

**<https://shs.hal.science/halshs-00096435>**

Submitted on 19 Sep 2006

**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.

# L'ANALYSE FORMELLE DES EGODOCUMENTS DANS UN SYSTÈME INFORMATIQUE DE PRODUCTION DE RESSOURCES ÉLECTRONIQUES

Dominique Taurisson  
CNRS / EHESS

## ***Corpus en lettres et sciences sociales : des documents numériques à l'interprétation***

Colloque international et école d'été, Albi, 10-14 juillet 2006, organisé dans le cadre des Colloques d'Albi Langages et signification (CALS)

[http://www.revue-texto.net/Archives/Albi\\_2006/Albi\\_programme.html](http://www.revue-texto.net/Archives/Albi_2006/Albi_programme.html)

## **SOMMAIRE**

1. Les egodocuments
2. Technologie de lecture
3. Contexte problématique
4. Instrumentation et analyse formelle
5. Analyse formelle et *Relateurs*
6. Instrumentation et production de ressources
  - 6.1. Croisement des sources et mutualisation
  - 6.2. Analyse formelle et projet éditorial
  - 6.3. Accessibilité des ressources
7. Bibliographie

**Résumé :** Les Egodocuments sont une mine d'informations pour les historiens et les littéraires, mais leur richesse et leur surabondance constituent souvent un obstacle à leur bonne exploitation. Certains systèmes d'analyse assistée permettent déjà d'en faire des explorations partielles ; nous présenterons, ici, une expérience d'analyse formelle assez différente, mais complémentaire.

Dans le cadre de travaux de production de ressources scientifiques électroniques utilisant un système informatique original et adapté (Arcane), les chercheurs rassemblent, organisent et analysent sources et documents de travail dans une base de données relationnelles considérée comme un « monde » : ils définissent les objets de ce monde (Sujets Personnes, Sujetslieux, SujetsInstitutions, SujetsOuvrages, etc), et formalisent les concepts et notions constitutifs des problématiques qui les intéressent sous forme de « relateurs » : par exemple, la rencontre, l'échange, le partage, le contrat, le déplacement, le transfert, etc. Les relateurs sont des expressions formelles composées d'un nom et d'une suite d'arguments : objets du monde, termes du métalangage, dates, modalités, etc. ; ils servent à établir des liens dynamiques entre les objets du monde et accroissent la combinatoire sémantique de ce monde.

Les chercheurs parcourent leurs documents textuels et images, et peuvent décomposer jusqu'à la plus petite unité de sens les informations « intéressantes de leur point de vue » que ces documents enferment. Ils les enregistrent et les stockent dans les

relations *ad hoc*, par instanciation de *relateurs*. Plusieurs relations peuvent décrire et être ancrées sur une même séquence de caractères, ou un segment d'image.

L'exposé montrera comment la mise en œuvre de ce mode électronique d'analyse, génère une qualité et une grande profondeur de lecture : textes et images se dévoilent au chercheur dans toute leur richesse, leur complexité intriquée, et leur feuilletage naturel (ce qui se voit et ce qui ne se voit pas), d'où une démultiplication du dialogue entretenu entre les documents et le chercheur.

Des utilisateurs de disciplines différentes peuvent d'ailleurs explorer au moyen d'autres *relateurs*, correspondant à leurs propres domaines d'enquête, les mêmes matériaux, et produire ainsi une multiplicité d'interprétations et de points de vue.

Les informations ainsi collectées et structurées peuvent être comparées, croisées, on peut en faire des analyses statistiques ou qualitatives et, également, en produire des représentations sous diverses formes graphiques. Ces informations, et les documents dont on les a extraites, sont exportables au format XML, à destination d'autres systèmes d'information.

Outre l'intérêt immédiat de faciliter les travaux personnels des chercheurs et de les initier à de nouveaux modes de production de la connaissance, ce système me semble associé de deux façons à la problématique du colloque :

- en permettant de développer un langage commun formalisé (objets, concepts) utilisable pour mettre en relation plusieurs « mondes » électroniques, et assurer communication et échanges entre eux, il repose, à nouveaux frais, la question de la mutualisation des recherches
- en mettant en œuvre de nouvelles formes d'élaboration et de diffusion des savoirs, il s'inscrit naturellement dans le débat sur les accès à ces savoirs

## 1. Les egodocuments

De nombreux chercheurs s'accordent aujourd'hui pour considérer que les egodocuments<sup>1</sup>, en tant que gisements exceptionnels d'informations, représentent une des voies d'accès possible à l'étude des relations interpersonnelles et des espaces relationnels des acteurs sociaux<sup>2</sup>, autrement dit, de la sociabilité comprise comme « l'ensemble des formes concrètes, des modalités, des structures et des processus de mise en communication et de socialisation des individus dans une société donnée »<sup>3</sup>.

Or, pour l'Age moderne, les correspondances, catégorie par excellence des egodocuments, sont encore assez régulièrement exclues des sources dites « du for privé »<sup>4</sup>. Elles ne sont pas toujours reconnues comme des documents sérieux et *vrais* du privé et donc fiables et exploitables pour la recherche, ni comme des sources neutres dès lors qu'elles déformeraient le « réel » puisqu'elles qu'elles obéissent à des règles de savoir-vivre et de mise en scène de soi par soi ».

Par ailleurs, traditionnellement absorbés par l'étude et la publication de corpus des personnages remarquables de cette époque, les chercheurs jugent souvent suspectes les correspondances produites par des individus modestes ou plus obscurs : des lacis de multiples arrangements, aux structures encore peu lisibles, où les agents sont tour à tour sollicités, recommandés, protecteurs, ceci au sein d'un réseau compliqué de relations de natures diverses aussi bien familiales, qu'administratives, commerciales ou politiques, souvent sans intérêt littéraire, parfois singulièrement dépourvues de pensée, à la fois considérables et lacunaires, truffées de mentions émietées relatives à la vie quotidienne (« monnaie courante » de la sociabilité), où l'on peine finalement à repérer des informations significatives.

Thomas Grosser soulève de son côté d'autres problèmes associés en général aux sources des historiens, et aux egodocuments en particulier : « De façon générale, on peut remarquer que la « banque de données » de l'historien est beaucoup plus *limitée et fragmentaire* que celle du sociologue ou du socio-psychologue. Car ces derniers peuvent analyser des phénomènes actuels à l'aide d'expériences et « créer » eux-mêmes leurs données. Au contraire, l'historien *ne peut pas* trouver pour chaque variable et chaque cas les données empiriques correspondantes à des relations de facteurs complexes, car il est dépendant d'informations qui lui sont transmises de façon fragmentaire et pas toujours dépourvues de *déformations*. Cependant, on peut *reconstruire une mosaïque* plus ou moins vraisemblable, qui permet d'éclairer les grandes lignes du processus et que les théories modernes permettent de modéliser<sup>5</sup>. » Outre l'intérêt de pointer précisément une question récurrente sur laquelle nous allons revenir, Grosser esquisse une des pistes de travail que nous avons scrupuleusement suivie : « la reconstruction de la mosaïque ».

## 2. Technologie de lecture

Pour réfléchir aux deux catégories fondamentales que T. Grosser dégage : données extraites et données créées, on peut se reporter à l'ouvrage de Maurizio Gribaudo, *Espaces, temporalités, stratifications. Exercices sur les réseaux sociaux*<sup>6</sup>, où un processus classique de création de données par des historiens est justement décrit.

L'équipe de Gribaudo a demandé à des personnes choisies de faire le compte rendu, dans des cahiers d'enregistrement déjà structurés, de leurs contacts quotidiens directs et indirects, de leur parcours biographique et de leurs modèles de référence relationnels.

<sup>1</sup> Néologisme proposé dans les années 1960 par le chercheur néerlandais Jacob Presser pour désigner les textes où l'on parle de soi.

<sup>2</sup> BEAUREPAIRE et TAURISSON, 2003.

<sup>3</sup> AGULHON, 1986.

<sup>4</sup> FOISIL, 1986.

<sup>5</sup> GROSSER, 1992.

<sup>6</sup> GRIBAUDI, 1988.

Les enquêtés détaillent, par exemple, leurs rencontres dans une période donnée : le moment de chaque rencontre, sa durée, son contenu, le lieu, le nombre et les personnes présentes, la description de la personne rencontrée, ainsi que les circonstances de leur première rencontre, etc. La collecte de données génère donc ici une production textuelle dont le contenu apparaît comparable à celui accumulé dans de nombreux egodocuments.

Dans ce cas (les cahiers d'enregistrement), la production informative peut être immédiatement formalisée : il y a structuration *a priori* de l'enquête, sous la forme d'une grille, d'un modèle, qui sert tout à la fois, à la réalisation de l'enquête elle-même, à l'analyse et à l'interprétation des résultats.

Dans l'autre cas, celui des chercheurs en sciences humaines confrontés à des egodocuments, il apparaît nécessaire d'élaborer, *a posteriori* cette fois, une sorte de technologie de lecture pour repérer et extirper des informations hétéroclites de leur gangue textuelle et les structurer selon des modèles adaptés.

Dans les deux cas, l'objectif est d'analyser les informations collectées ou « dénichées », la place et le rôle des agents qu'elles concernent, qu'il s'agisse de personnes, de groupes, de lieux ou de productions, ainsi que les mécanismes de leurs interrelations.

Pour abonder dans le parallèle établi entre données créées et données extraites, l'équipe de Gribaudi, elle-même, montre que les données créées sont à prendre avec autant de prudence que les données des historiens. Les participants d'une enquête ont, en effet, des réactions diverses qui peuvent aller jusqu'à la parasiter, certains s'en servent comme d'une source d'inspiration, d'autres sont des « répondants à problème » qui accumulent retards, oublis, réticences, négligences et qu'il faut accompagner pendant des mois pour les aider à gérer la rédaction de leur cahier... Par ailleurs, rien n'assure que les données récoltées soient complètes, objectives, et *vraies*, même dans le cas des « répondants sans problèmes », le facteur psychologique n'étant pas vraiment quantifiable. M. Gribaudi écrit d'ailleurs qu'on obtient finalement : « trop de nuances, trop de renseignements, trop de vies individuelles... ».

Même si l'on convient unanimement de la nécessité d'élaborer une technologie de lecture adaptée capable d'aider les chercheurs à réaliser des micro-analyses de leurs sources, il ne faut pas perdre de vue que s'il est fort *probable* que les informations contenues dans des egodocuments peuvent leur apprendre quelque chose des comportements de ceux qui les ont rédigés, il est, en revanche, fort improbable de découvrir dans *un* corpus (la correspondance de X, par exemple) toutes les données significatives concernant le scripteur : notamment toutes les personnes qu'il connaît, tous les lieux auxquels il est lié. *A fortiori*, les liens suggérés par la liste de ses correspondants et par le contenu des lettres échangées ne constituent certainement pas l'ensemble des relations de cette personne, ce qui fait qu'il n'est guère possible d'expliquer les stratégies individuelles par les seuls témoignages des individus concernés.

L'abondance des documents et des renseignements qu'ils contiennent se conjugue donc paradoxalement à « un manque », à une absence, que l'instrumentation informatique mise au point devra compenser par la possibilité donnée aux chercheurs de croiser et de confronter différents corpus et sources documentaires.

### **3. Contexte problématique**

La plupart des informations enfouies dans les egodocuments de l'Age moderne parvenus jusqu'à nous concernent des *SUJETS*, personnes, lieux, institutions, groupes (familles, organisations), productions écrites, artistiques, idées, et des *LIENS directs ou indirects* tissés entre ces Sujets.

Or, pour circonscrire l'espace relationnel d'un Sujet, notamment dans le cadre de travaux sur la sociabilité, il est nécessaire de savoir pour une personne, qui et qu'est ce qu'elle connaît, quels usages elle fait de ces connaissances, quels profits elle en retire,

comment elle les a acquises, comment elle les entretient à travers les différents états de sa vie ; pour un lieu ou une institution, qui a eu des contacts directs ou indirects avec ce lieu (qui l'a visité ou habité), quand, pourquoi ; pour un groupe, qui en fait partie, depuis quand, à quelles opérations ses membres ont-ils été associés ; pour une production, qui l'a créée, qui l'a lu ou vu, dans quel endroit, qui en a parlé, etc.

Ce qui équivaut pour un Sujet à un ensemble variable de liens évolutifs inscrits dans différents contextes, et donc à un espace relationnel sans cesse reconfiguré.

En outre, chaque lien est à étudier en fonction de tous les autres liens, de la position relationnelle et des « attributs » (propriétés, connaissances) des Sujets, ces trois facteurs s'influençant les uns les autres, d'où évidemment une forte interaction entre identités, comportements et espaces relationnels (espace relationnel du Sujet et espaces relationnels des autres Sujets en relation avec lui). : dans le cas des relations interpersonnelles, interviennent donc les caractéristiques de chaque individu concerné, les contacts qu'ils entretiennent, ainsi que les règles ou codes de comportement en vigueur à un moment précis. Autrement dit, s'intéresser à leurs parcours signifie étudier en même temps les espaces relationnels dans lesquels ils s'inscrivent, tout en rendant compte de la structuration de ces espaces.

Ce qui revient à définir l'espace relationnel d'un individu à un instant « t » comme la concrétisation momentanée de son parcours à l'intérieur de l'espace social, son « capital relationnel » (humain et social)<sup>7</sup>, et son espace relationnel global comme la somme de ses identités, de ses trajectoires et des liens enregistrés à tous les instants « t » de sa vie.

Si l'espace relationnel complet d'une personne paraît dès lors singulièrement inaccessible, il semble, en revanche, possible d'isoler et d'observer certains des segments qu'elle mobilise dans le cadre de stratégies personnelles à des moments précis de sa vie, et de comparer les circonstances et les modes de leur activation, leur évolution dans le temps, éventuellement d'étendre, au travers d'une démarche comparative, ces observations à d'autres corpus et à d'autres personnages.

Par extrapolation, on pourrait tenter de cerner l'intérêt d'un Sujet Lieu en circonscrivant son « espace relationnel » à différentes périodes, en notant les configurations de personnes et de pratiques (activités) associées à ce lieu : les rapports périodiques ou ponctuels entretenus par différents types de Sujets avec ce lieu, et leurs transformations, les personnes qui s'y sont rencontrées, qui en ont parlé, qui y sont revenu, les œuvres qui y ont été créées, celles qui ont pris ce lieu comme espace de narration, etc. Le travail consisterait à élaborer une typologie des lieux à partir de l'étude des fonctions relationnelles dont ils sont habituellement le théâtre, et dont on établirait une nomenclature

Dans le même esprit, et pour revenir aux individus et aux groupes, il s'agirait de se concentrer sur leurs comportements et leurs agissements vus à travers quelques grands types de liens relationnels, assimilés à des formes élémentaires de la sociabilité : la recommandation, la rencontre, le partage, l'échange sous ses différentes modalités (don, prêt, achat-vente, etc), méthode qui permettrait d'agréger et d'observer des Sujets d'une façon moins egocentrée que dans les études de réseaux traditionnelles.

Suivre ces différentes pistes rapidement esquissées, revient à *préconiser* pour l'analyse des egodocuments, notamment des correspondances, des cheminements scientifiques qualitatifs et compréhensifs, et pas seulement quantitatifs, de manière – à améliorer les modèles destinés à instancier et à expliquer, à partir de la multiplicité des pratiques et des micro-histoires, les formes et les invariants des comportements sociaux, et – à vérifier l'hypothèse que c'est de « l'ensemble des choix individuels que résultent certains processus macroscopiques ». Et bien entendu, nous souscrivons pleinement à ce qu'Alain Bensa rajoute, à savoir qu'il faut évidemment se doter *des outils et des moyens méthodologiques* pour étudier dans les textes les événements microhistoriques singuliers que nous venons de décrire, et pour les relier à d'éventuels systèmes plus englobants de données et de significations<sup>8</sup>.

---

<sup>7</sup> LEMERCIER, 2005.

<sup>8</sup> BENSA, 1996, p. 40.

#### 4. Instrumentation et analyse formelle

Nous cherchons donc à déceler ce qui se voit et ce qui ne se voit pas, ou mal, ce qui n'est pas explicite, à percevoir et à décomposer les feuilletages et les intrications d'information, puis à isoler et réorganiser les unités de connaissance ainsi récoltées pour multiplier les angles de vue et les échelles d'observation ; enfin, à offrir à différents types de lecteurs des accès pratiques pour mener leurs propres recherches : une façon de laisser parler les textes et de mettre en ordre la grande masse des éléments dispersés, combinés et hétéroclites qu'ils contiennent<sup>9</sup>.

Pour toutes ces raisons, et dans l'objectif à moyen terme de favoriser les investigations croisées dans les egodocuments, on a conçu un instrument électronique (ArcaneWriter) et proposé un modèle de structuration de la connaissance (Arcane)<sup>10</sup>. L'un et l'autre associés, permettent aux chercheurs d'enregistrer et d'étudier, sous forme de bases de données propres, mais connectables, des documents multimédias (en particulier des egodocuments textuels) dans leur transcription philologique et dans leur matérialité initiale (images des manuscrits) : au fil d'une lecture approfondie, le chercheur enrichit ses documents au moyen d'une indexation systématique sous forme de notions-concepts très finement spécifiées ; dans le cas des correspondances de l'époque moderne et d'études sur la sociabilité : la circulation, la transmission, l'appartenance, l'échange, le transfert, ou l'aide, ... (autant de signes de relations entre personnes, lieux, institutions, idées, productions, et objets) : il note tout ce qui peut caractériser et relier les individus, les lieux, les groupes, les idées et les productions, ainsi que toute indication relative à la datation, à l'intensité et à la durée des relations enregistrées.

Nous verrons qu'en rendant ainsi possible la manipulation simultanée de différents contextes, échelles et points d'approche, un outil d'observation-enregistrement comme Arcane, sorte de microscope pour molécules textuelles, peut doublement aider le chercheur – à construire des visions formalisées et des répliques même fragmentées des modes de vie et des environnements des acteurs sociaux, – dans le cadre d'études sur les interrelations, à envisager la mise au jour d'une sorte de typologie des « formes de la sociabilité » considérées non plus comme des processus singuliers et isolés dans une époque et un espace donnés, mais comme des phénomènes stables que l'on découvre inexorablement à l'œuvre, selon des modalités à étudier, dans différentes formes de sociétés. L'hypothèse étant bien que ces catégories relationnelles sont génériques, peut-être invariantes, et qu'elles participent, il s'agit encore de comprendre dans quelle mesure, « à la construction et à la fortification des liens sociaux ». En 1986, Maurice Agulhon avait déjà proposé de choisir un cadre d'étude circonscrit et d'étudier l'ensemble *des formes de sociabilité*, qu'elles soient institutionnelles ou privées, à un moment donné pour mieux en percevoir les continuités et les interférences<sup>11</sup>, et il avait été l'un des premiers, dès les années 60<sup>12</sup>, à travailler sur le comportement privé et quotidien, « susceptible d'analyses relationnelles », alors même qu'il ne disposait d'aucun outil informatique pour mener à bien cette tâche.

#### 5. Analyse formelle et *Relateurs*

Le principe est donc simple : chaque chercheur ou équipe de recherche nourrit sa propre base de données structurée compatible avec l'architecture *Arcane*, tout en conservant sa logique scientifique, ses problématiques de recherche, et ses contraintes éditoriales. Il accepte les principes et modalités de rassemblement de l'information préconisés, et en conséquence, de mettre à disposition du public et de ses collègues

<sup>9</sup> *Alter histoire*, 1991, p. 182.

<sup>10</sup> LOCHARD et TAURISSON 2000 ; LOCHARD et TAURISSON 2001.

<sup>11</sup> AGULHON, 1986.

<sup>12</sup> AGULHON, 1963.

chercheurs un certain nombre des données signées, qu'il a rassemblées et traitées.

Quant à la mise en œuvre de la méthode d'analyse formelle des textes exposée succinctement plus haut, et définie dans le paradigme *Arcane*, elle peut être décrite et caractérisée par quatre opérations principales faciles à adapter aux problématiques particulières de chaque chercheur :

- Le chercheur définit, en rapport avec ses objets de recherche, les notions-concepts qu'il se propose de repérer dans les documents composant son corpus. Il isole les unités sémantiques qui constituent ces notions-concepts de manière à pouvoir les formaliser et les implémenter dans un système d'information électronique sous forme de *relateurs*.

Les *relateurs* sont des expressions « prédicatives » formées par une suite de valeurs ordonnées qui permettent d'établir des liens entre les Sujets (Lettres, Personnes, Lieux, Produits, Idées, Dates, Événements) et des expressions métalinguistiques, selon des combinaisons formelles.

- Le chercheur sélectionne et enrichit systématiquement des séquences de caractères ou d'image, en les liant à un ou plusieurs de ces *relateurs*.
- La quantité considérable d'informations élémentaires (unités de sens) ainsi collectées et enregistrées peuvent elles-mêmes, selon des mécanismes intellectuels et électroniques, être mises en relation, recombinaisons et interprétées selon des problématiques claires et propres à chaque utilisateur. Ces mécanismes simples permettent de développer de nombreux autres traitements, et d'enrichir la sémantique du « monde » commun au chercheur et à son lecteur en définissant des requêtes pour l'interroger. L'exécution dynamique de ces requêtes provoque la sélection et l'affichage des objets vérifiant ces requêtes. Avec ces résultats, il est possible d'éditer des listes et de réaliser de nombreuses représentations graphiques dynamiques *in situ* dans l'environnement *Arcane*, ou par export vers d'autres applications mieux adaptées aux demandes spécifiques des chercheurs. Un système comme *Arcane* est aussi bien équipé des outils nécessaires pour représenter les connaissances, et les textes d'où elles ont été extraites, que pour les éditer simultanément sur supports électronique (*via XML*) et papier (*via TeX*), nous en reparlerons plus loin.

## 6. Instrumentation et production de ressources

### 6.1. Croisement des sources et mutualisation

Il est évident que la manipulation d'un tel système allait naturellement conduire ses utilisateurs à vouloir étendre leurs recherches à d'autres corpus et sources d'information, à entreprendre des échanges et des croisements, opérations difficiles à mener jusqu'à maintenant, autrement que manuellement ou par de lourdes recherches automatisées, compte tenu des outils électroniques disponibles et de la difficulté de mettre en œuvre des travaux collectifs concertés et interopérables. Cette perspective devenue aujourd'hui réaliste constitue de notre point de vue, nous l'avons déjà souligné, une condition incontournable à un travail sérieux mené à partir de sources egodocumentaires, en particulier à cause de leur manque intrinsèque de fiabilité.

Plusieurs chercheurs et équipes expérimentent déjà la méthode d'analyse textuelle décrite plus haut<sup>13</sup>, d'où la nécessité de maintenir une cohérence sémantique dans le choix des concepts manipulés et dans leur formalisation. L'objectif étant d'élaborer des méthodes *générales* d'analyse formelle des contenus textuels adaptées, en particulier, à l'étude des relations interpersonnelles et plus généralement des processus inscrits dans les egodocuments, il s'agit d'établir collectivement un ensemble minimum commun de *relateurs* utilisables dans les différents projets éditoriaux et/ou de recherche des

<sup>13</sup> Voir TAURISSON, 2005, *Actes des Journées d'étude sur l'instrumentation Arcane*.

utilisateurs ; ces *relateurs* constituant eux-mêmes un outil de production de ressources partageables.

Dans l'idéal, il serait souhaitable que quels que soient ces projets, et en dépit de leur diversité, il soit possible aux différentes bases de données qui les portent de « communiquer » de façon rationnelle et dynamique : or, pour établir des bases de connaissances disponibles et accessibles à tous, capables de dialoguer entre elles, il faut s'entendre sur ce qu'on enregistre et sur la manière de l'enregistrer. C'est pourquoi un intense effort de normalisation est mené au sein du groupe Arcane pour définir un métalangage commun (une grammaire) pour structurer, baliser, nommer, etc. les données. Cela veut dire, notamment, élaborer collectivement des listes de types de Sujets avec leurs descripteurs (personnes, lieux, institutions, thèmes, productions...), de *relateurs* adaptés à l'étude de problématiques voisines ou convergentes (voir plus haut), de genres de documents, d'identificateurs universels, et d'énumérations descriptives<sup>14</sup>.

En ce qui concerne les recherches sur les formes de sociabilité, on entrevoit déjà bien les bénéfices importants qu'elles pourraient retirer de la mutualisation et du croisement des sources, ne serait-ce que, nous l'avons déjà dit, pour compléter et confronter les informations manipulées. Quant à l'étude de l'instanciation des concepts fondamentaux qui relèvent de la sociabilité, comme la rencontre ou l'aide, elle trouverait les plus grands avantages dans la mise à disposition des chercheurs de très nombreuses occurrences comparables collectées dans des corpus de différentes périodes, et dans plusieurs types de documentation : contextes des rencontres, contraintes, processus, déroulements, résultats, etc. ; contextes des aides reçues, demandées, négociées, remerciées, modes de sollicitation, conséquences, effets, évolution des acteurs, etc.

## 6.2. Analyse formelle et projet éditorial

Ces opérations de mise en réseau de travaux individuels et de normalisation des méthodes d'enquêtes dans les sources textuelles vont naturellement de pair avec la volonté d'associer étroitement l'analyse de ressources documentaires multimédias à leur diffusion dans la communauté scientifique (éditions sous multiformats) et à la production simultanée de résultats de recherche ; il ne suffisait donc pas de mettre à la disposition des chercheurs une instrumentation *ad hoc* fortement adossée à leurs pratiques de recherche, mais il fallait l'adapter aux mécanismes qui ponctuent traditionnellement le circuit de l'édition : autrement dit, il fallait inventer un double outil d'écriture et de publication.

Le chercheur devait pouvoir mettre en œuvre, au sein de sa propre base de données et de façon autonome, différentes entreprises allant de l'écriture, à la recherche pure, jusqu'à la publication multisupports de documents de base (textes, images fixes et animées, son) et de connaissances scientifiques (papier, Web, CD-ROM), opérations respectant évidemment toutes les règles éditoriales préconisées par la communauté scientifique (notes de commentaire, notes critiques, mise en page élaborée, index et bibliographie cumulatifs, etc.).

Tout cela a été rendu possible par la mise au point d'un paradigme qui articule le couple auteur-lecteur autour d'un *monde* de connaissances structuré que l'auteur élabore et que le lecteur parcourt pour le comprendre : un ensemble réticulé de connaissances scientifiquement validées et architecturées dans lequel, grâce à un métalangage commun, les auteurs et les lecteurs peuvent agir collectivement<sup>15</sup>, l'édition au sens traditionnel ne constituant plus qu'une des phases d'un processus coopératif sans limites dans le temps et dans l'espace qu'on pourrait détailler ainsi :

<sup>14</sup> Voir TAURISSON, 2005, *Actes des Journées d'étude sur l'instrumentation Arcane*.

<sup>15</sup> LOCHARD, 2005, « Les documents électroniques dans le système d'écriture et d'édition Arcane » <<http://hal.ccsd.cnrs.fr/ccsd-00016064>>.

- Définir l'architecture du monde à éditer comme instance d'une méta-architecture
- Construire le monde des Sujets d'intérêt du projet éditorial étroitement associé au programme de recherche
- Enrichir ce monde par des documents multimédias et des relations (occurrences des *relateurs*) pour le décrire, le structurer, l'illustrer, l'interpréter et le représenter
- Enfin, en extraire à des moments choisis des sous-ensembles cohérents pour produire et publier des livres électroniques.<sup>16</sup>

### 6.3. Accessibilité des ressources

L'accessibilité des ressources est aujourd'hui un problème central pour la diffusion de la connaissance par internet et de nombreux travaux y sont consacrés.

Au contraire des heuristiques qui favorisent des traitements en aval, au moment de la publication, traitements lexicologiques et linguistiques notamment, des outils d'écriture comme Arcane cherchent à donner aux auteurs les moyens de penser, de concevoir et de réaliser des accès intentionnels aux ressources qu'ils produisent, de manière à les cartographier conceptuellement, à en repérer les points d'entrée, et à baliser à l'intention de leurs différents types de lecteurs des navigations logiques dans l'hypergraphe des données ainsi constitué. Les *relateurs* jouent un rôle particulièrement important dans l'élaboration de cette sémantique de communication.

## 8. Bibliographie

AGULHON, Maurice, 1963, « Les Associations, confréries religieuses et loges maçonniques en Provence orientale à la fin de l'Ancien Régime », *Actes du Congrès national des Sociétés savantes* ; section d'histoire moderne et contemporaine 1963, 87, p. 73-86.

AGULHON, Maurice, 1986, « La sociabilité est-elle objet d'histoire », *Sociabilité et société bourgeoise en France, en Allemagne et en Suisse, 1750-1850*, sous la dir. d'E. François, Paris, Recherche sur les Civilisations, 1986, p. 13-22 (*Travaux et Mémoires de la Mission historique française en Allemagne*).

*Alter histoire : essais d'histoire expérimentale*, 1991, Daniel S. Milo et Alain Boureau avec Hervé Le Bras, Paul-André Rosental [et al], Paris, Les Belles Lettres, 1991 (*Histoire*).

BEAUREPAIRE, Pierre-Yves et TAURISSON, Dominique, (éds.), 2003, *Les ego-documents à l'heure électronique : nouvelles approches des espaces et des réseaux relationnels*, Actes du colloque tenu du 23 au 25 octobre 2002 (CNRS & Université de Montpellier), Montpellier, SerPub, 2003, 554 p. (édition en libre consultation sur le Web : <<http://www.univ-montp3.fr/arcanews/egodoc/>>).

BENSA, Alain, 1996, « De la micro-histoire vers une anthropologie critique », *Jeux d'échelle. La micro-analyse à l'expérience*, dir. Jacques Revel, Paris, Gallimard, Le Seuil, 1996, p. 37-70.

FOISIL, Madelaine, 1986, « L'écriture du for privé », *Histoire de la vie privée* sous la dir. de P. Ariès et G. Duby, Paris, Seuil, tome III (*De la Renaissance aux Lumières*), 1986, p. 331-369.

<sup>16</sup> Voir sur le Web l'édition du *Journal de Corberon (1775-1780)* réalisée avec Arcane dans les conditions décrites : *Web passif* : <http://egodoc.revues.org>, et *Web actif* : <http://www.univ-montp3.fr/arcanews/egodoc/> ; voir aussi TAURISSON, 2004.

GRIBAUDI, Maurizio, 1998, *Espaces, temporalités, stratifications. Exercices sur les réseaux sociaux*, Paris, EHESS, 1998.

GROSSER, Thomas, 1992, « Les voyageurs allemands en France. Etudes de cas et perspectives d'analyse », *Allemands en France, Français en Allemagne 1715-1789. Contacts institutionnels, groupes sociaux, lieux d'échanges*, dir. par Jean Mondot, Jean-Marie-Valentin et Jürgen Voss, 1992, p. 209-235.

LEMERCIER, Claire, 2005, « Analyse de réseaux et histoire », *Revue d'histoire moderne et contemporaine*, 52-2, avril-juin 2005, p. 88-112.

LOCHARD, Eric-Olivier et TAURISSON, Dominique, 2000, « The World according to Arcane : an operating instrumental paradigm for scholarly edition », *Proceedings of the international conference organized by the Constantijn Huygens Instituut and the Free University Amsterdam, Den Haag, 7-8 décembre 2000*, Berlin, Weidler Buchverlag, 2002, p. 151-162.

LOCHARD, Eric-Olivier et TAURISSON, Dominique, 2001, « Le monde selon Arcane », *Le Document au XXI<sup>e</sup> siècle, Cahiers Gutenberg*, n° 39-40, mai 2001, p. 89-105.  
<<http://www.gutenberg.eu.org/publications/cahiers/r25-cahiers39-40/129-lochard.html>>

LOCHARD, Eric-Olivier et TAURISSON, Dominique, 2002, « Correspondances, réseaux, édition électronique », *La Plume et la Toile. Pouvoirs et réseaux de correspondance dans l'Europe des Lumières*, textes réunis par Pierre-Yves Beaurepaire, Arras, Artois Université Presses, 2002, p. 171-192.

LOCHARD, Eric-Olivier, 2005, « Les documents électroniques dans le système d'écriture et d'édition Arcane » <<http://hal.ccsd.cnrs.fr/ccsd-00016064>>

TAURISSON, Dominique, et Pierre-Yves Beaurepaire, (éds.), 2000 -->, Edition électronique du *Journal (1775-1781)* de Marie Daniel Bourrée, chevalier de Corberon <<http://www.univ-montp3.fr/arcanews/egodoc/>>.

TAURISSON, Dominique, 2004, « *Le journal de Corberon* sur le Web : édition numérique ou édition électronique ? », Actes du colloque *La numérisation des textes et des images : techniques et réalisations*, Université de Lille 3, 16 et 17 janvier 2003, textes réunis et édités par Isabelle Westeel et Martine Aubry, CeGes-Lille 3, 2004., p. 47-62.

TAURISSON, Dominique, éd., 2005, *Actes des Journées d'étude sur l'instrumentation Arcane*, organisées les 17 et 18 novembre 2005 par J. Boutier et D. Taurisson, SHADYC (EHESS-CNRS, UMR 8562). <<http://www.egodoc.revues.org/journeesArcane/>>