

Schriften des Instituts für Dokumentologie und Editorik – Band 3

Kodikologie und Paläographie im digitalen Zeitalter 2

Codicology and Palaeography in the Digital Age 2

herausgegeben von | edited by

Franz Fischer, Christiane Fritze, Georg Vogeler

unter Mitarbeit von | in collaboration with

Bernhard Assmann, Malte Rehbein, Patrick Sahle

2010

BoD, Norderstedt

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de/> abrufbar.

© 2011

Online-Fassung

Herstellung und Verlag der Druckfassung: Books on Demand GmbH, Norderstedt 2010

ISBN: 978-3-8423-5032-8

Einbandgestaltung: Johanna Puhl, basierend auf dem Entwurf von Katharina Weber

Satz: Stefanie Mayer und L^AT_EX

Archives d'un lecteur philosophe. Le traitement numérique des notes de lecture de Michel Foucault

Samantha Saïdi, Jean-François Bert, Philippe Artières

Résumé

Nous détaillerons le travail effectué sur une archive de pièces manuscrites : essentiellement des fiches de lecture, des fiches thématiques et des fiches bibliographiques, que le philosophe Michel Foucault rédigea et rassembla en un instrument de travail notionnel pour la rédaction de son essai, *Les mots et les choses*. Cet échantillon de sa bibliothèque de travail privée constitue un véritable champ d'investigation de la façon dont Michel Foucault utilisait ses sources primaires et secondaires : comment il les lisait, les transcrivait et, enfin, les restituait dans ses ouvrages. Bien que ce travail d'investigation ait pour matériau principal un objet qui l'exclut a priori d'une approche codicologique (codex, parchemins vs papier du XXème s.) ou paléographique (époque), nos méthodes sont proches et nous voudrions partager ici avec les spécialistes de ces disciplines, notre approche propre à l'anthropologie des pratiques culturelles, mais surtout notre expérience de manipulation des outils et méthodes de l'archivistique, de l'informatique et du numérique. En effet, de la Description Archivistique Encodée (EAD) au Langage de Modélisation Unifiée (UML), ces outils et méthodes méritent d'être décrits car c'est grâce à eux que ce corpus qui donne à voir la « bibliothèque » choisie par le philosophe pour la rédaction des *Mots et les choses*, est aujourd'hui intelligible et surtout utilisable pour la recherche.

Zusammenfassung

Dieser Beitrag erläutert, auf welche Weise ein Team aus Anthropologen und Informatikwissenschaftlern die handschriftlichen Vorarbeiten Michel Foucaults zu *Les mots et les choses* (dt.: *Die Ordnung der Dinge*) erschlossen haben: die Digitalisierung eines Konvoluts aus Exzerpten, thematischen Aufzeichnungen und bibliographischen Notizen. Es handelt sich bei dieser privaten »Arbeitsbibliothek« um eine ideale Quellenbasis, um Aufschlüsse über Foucaults Schaffensprozess und seinen Umgang mit Primär- und Sekundärquellen zu erlangen: Wie hat Foucault relevante Texte gelesen, transkribiert, exzerpiert und in seinem Werk weiterverwendet? Obgleich es sich um zeitgenössische Notizen auf Papier und somit nicht um einen Gegenstand kodikologischer oder paläographischer Forschung im engeren Sinne handelt, ist der methodische Zugriff derselbe oder

doch zumindest weitgehend übertragbar und soll aus diesem Grund in diesem Band zur Diskussion gestellt werden. Mögen die Ausgangsfragen einem anthropologischen Interesse an kulturellen Praktiken entspringen, so geht es bei der Quellenerschließung in erster Linie um die Verwendung digitaler archivalischer Hilfsmittel und Methoden. Von der Verwendung des XML-Standards EAD (Encoded Archival Description) zur vereinheitlichten Modellierungssprache UML (Unified Modeling Language) werden die Methoden vorgestellt, dank derer Foucaults Auswahlbibliothek zur Vorbereitung von *Les mots et les choses* in einem virtuellen Archiv sowohl der Forschung als auch der interessierten Öffentlichkeit zugänglich gemacht werden konnten.

Abstract

In this chapter we explain how our team of anthropologists and information scientists worked on Michel Foucault's characteristic set of handwritten material: mainly reading cards, bibliographical and subject cards that Michel Foucault wrote, filed and used before writing his essay *Les mots et les choses* (*The Order of Things*). This digitised sample of his private "working library" constitutes an exceptional medium to investigate, discover and analyze how Foucault used primary and secondary sources for his writing: how he read them, extracted from them, adapted them in his handwritten transcriptions and in the end transformed these transcriptions into his final work. Even though this handwritten material is contemporary and thus does not convey an object for codicological or palaeographical research in a strict sense, our perspective and the methodology and techniques applied are similar. This sample was indeed an ideal corpus for testing the use of digital and data-processing methods and tools. By these means we were able to keep a digitised record of the archive, by describing, annotating and extracting indexes, but they also gave us the means of modeling in UML the final web application that will bring together the EAD based catalogue and the virtual Foucault private library. The portal will enable researchers and, to some extent, the general public to see, understand and use the "library" chosen by the philosopher for the drafting of *Les Mots et les choses*.

Introduction

Dans un texte intitulé « À propos des faiseurs d'histoire », autour d'une affaire de plagiat qui touche un livre de Jacques Attali, Michel Foucault indique :

[C]e qui me paraît indispensable, c'est le respect à l'égard du lecteur. Un travail doit dire et montrer comment il est fait. C'est à cette condition qu'il peut non

seulement ne pas être trompeur, mais être positivement utile. Tout livre dessine autour de lui un champ de travail virtuel et il est jusqu'à un certain point responsable de ce qu'il rend possible ou impossible. Un livre – je parle bien sûr de ces ouvrages de savoir – qui brouille ses manières de faire n'est pas quelque chose de très bien. Je rêve de livres qui seraient assez clairs sur leur propre manière de faire pour que d'autres puissent s'en servir librement, mais sans chercher non plus à brouiller les sources. La liberté d'usage et la transparence technique sont liées.

(*Dits et écrits* 414)

Quelques années auparavant, à Raymond Bellour, il souligne le fait que, dans sa méthode d'analyse des documents, tout devient source historique : « Il ne doit pas y avoir de choix privilégié. Il faut pouvoir tout lire, connaître toutes les institutions et toutes les pratiques. [...] Ce qui fait qu'on traitera dans la même foulée *Don Quichotte*, Descartes et un décret sur la création des maisons d'internement par Pomponne de Bellièvre » (Bellour 17).

En regard de cette manière de présenter ses propres recherches, la possibilité est aujourd'hui offerte de reprendre cette question des pratiques réelles de Foucault à partir des nouvelles méthodologies de recherche et d'enquête qui se sont développées autour des archives du travail intellectuel et qui ont été insufflées par Bruno Latour, Roger Chartier (cf Kraus) et, plus récemment, Christian Jacob et sa notion de pratique savante qui lui permet de regrouper sous cette notion opératoire l'ensemble des opérations manuelles, discursives et intellectuelles mobilisées dans la production ou la réception d'un savoir.

Dans son célèbre article « Les Vues de l'esprit, une introduction à l'anthropologie des sciences et des techniques », et à propos des laboratoires scientifiques Bruno Latour donne en détail le programme de cette anthropologie du savoir dans laquelle s'inscrit cette recherche :

Pourquoi les inscriptions de toutes sortes sont-elles aussi importantes pour les chercheurs, les ingénieurs, les architectes, tous ceux qui pensent avec leurs yeux et leurs mains ? Parce qu'elles offrent un avantage unique lors des discussions : « Vous doutez de ce que je dis ? Vous allez voir, je vais vous montrer ! » et sans remuer de plus de quelques centimètres, l'orateur déploie devant les yeux de ses critiques autant de figures, diagrammes, planches, silhouettes qu'il en faudra pour convaincre. Aussi médiates que soient ces inscriptions, aussi lointaines que soient les choses dont on parle, des chemins à double voie s'établissent. L'objecteur se trouve dominé par le nombre de choses dont parle l'orateur, toutes présentes dans la salle. Il peut douter de chacune d'elles, mais toutes ensemble, elles composent une formidable preuve. Nous sommes tellement

habitué à recourir à ces alliés ; que nous avons oublié ce que c'est que penser sans index, sans bibliographies, sans dictionnaires, sans fiches bristol, sans physiographes, sans cartes, sans diagrammes [...]. (87)

Plus récemment, l'historien Christian Jacob a voulu, en insistant sur la matérialité du travail intellectuel, redonner une nouvelle direction à ce même programme :

[S]i un anthropologue pénètre dans le bureau d'un universitaire contemporain ou dans la salle d'une soutenance de thèse, ou encore dans une bibliothèque ou une salle de colloque, sans préjuger du contenu et du sens des activités qui y prennent place, il sera amené à observer des gestes, des déplacements, des postures, une répartition spatiale des acteurs, un mobilier, des ordinateurs, des étagères avec des livres dont la répartition est signifiante, etc. S'il se penche sur le bureau d'un chercheur, il remarquera des pages imprimées, des fiches, un livre ouvert, différents crayons et stylos, une agrafeuse, un écran et un clavier d'ordinateur, des post-it, [...] des croquis, des paragraphes de texte, une carte, un tableau, des statistiques. [...] Notre hypothèse de travail est que ces différentes échelles d'observation, de l'environnement architectural à l'écran d'ordinateur, nous apprennent quelque chose de signifiant sur la nature du travail savant, au sens de séquences de gestes et d'opérations où le mental est relié au matériel, par le biais d'instruments, d'objets et d'inscriptions.

(Müller 122)

Il y a tout intérêt à ouvrir un terrain d'enquête sur les traces de Michel Foucault en bibliothèque qui informerait, de manière inédite, sur les procédures de lecture qui étaient les siennes. Autrement dit, analyser ces gestes minimes, répétés quotidiennement, qui font le métier d'intellectuel dans la deuxième moitié du XXe siècle. Grâce à Daniel Defert, l'on sait que la chaîne d'écriture de Foucault était composée d'au moins trois moments : « d'abord la version de ce qu'il ne fallait pas dire, qu'on pensait un peu spontanément. Ensuite, il y avait la reprise de tout ça à partir d'un travail de recherches, ce qui demandait facilement trois ans. Une fois ce travail de recherches fait, il y avait une réécriture » (Bellon 4). Dans cette enquête sur les « inscriptions » multiples de son activité de philosophe, le cas des *Mots et les choses* qu'il publia dans la collection de la « Bibliothèque des sciences humaines » chez Gallimard en 1966 est exemplaire de sa pratique.

En effet, le dossier préparatoire des *Mots et des Choses* que Daniel Defert¹ a mis temporairement à notre disposition, révèle que Michel Foucault rédigea en amont plusieurs centaines de fiches de prises de notes. Ces fiches témoignent à la fois chez

¹ Qu'il nous soit permis de remercier ici Daniel Defert de nous avoir offert la possibilité de consulter ces documents.

lui, d'une manière de cerner un domaine (comme la grammaire générale, l'histoire naturelle, l'économie...), de baliser un champ de recherche et d'organiser l'ordre du travail. L'existence de ces fiches n'est pas sans rapport non plus avec la méthode « archéologique » que Foucault met alors en place et qui doit lui permettre d'identifier les concepts par les connexions qui régissent leur emploi. Cette analyse, ajoute-t-il, a pour but de « reconstituer le système général de pensée dont le réseau, en sa positivité, rend possible un jeu d'opinions simultanées et apparemment contradictoires » (*Les Mots et les choses* 89–90).

Bien que notre approche se rapproche de l'anthropologie des pratiques culturelles et savantes et que ce travail d'investigation ait pour matériau principal un objet qui l'exclut à priori – si on s'en tient à des définitions strictes – d'une approche codicologique (codex, parchemins vs papier du XX^{ème} s.) ou paléographique (époque), nous avons voulu conserver, au vu de la nature du dossier, certaines techniques d'analyse matérielle propre à l'analyse des manuscrits modernes (Bustarret 47–60) mais surtout des techniques de l'archivistique pour reconstituer certaines habitudes d'écriture et de travail du philosophe ; comme, par exemple, sa manière d'assembler ses matériaux en liasses ou, au contraire, de les laisser sous forme de feuilles volantes.

Nous avons entrepris l'analyse matérielle de cet ensemble de documents, de ses découpages (dossiers, sous-dossiers, listes bibliographiques, fiches,...), ainsi qu'une description systématique de chacun des feuillets (rature, soulignement, organisation de la fiche) en procédant à un inventaire XML, qui, une fois complet, renseignera sur les usages d'une méthode de travail explicite. Par ailleurs, ce travail d'inventaire nous a obligé à mieux renseigner les circonstances de la transmission de ces documents et à analyser s'ils avaient fait, ou non, l'objet d'interventions ultérieures.

Au delà de l'approche disciplinaire, l'objet de cet article sera de partager avec les spécialistes de la codicologie et de la paléographie, des outils et méthodes du numérique qui nous ont permis de :

- numériser l'ensemble des documents manuscrits qui composent ce dossier,
- créer un catalogue informatisé de notices descriptives de ces fiches
- et, enfin, de concevoir une véritable bibliothèque virtuelle qui réunira à la fois ce catalogue numérique et l'ensemble des fiches numérisées.

Du protocole de numérisation au Langage de Modélisation Unifiée (Unified Markup Language ou UML) pour modéliser le site web final, en passant par la Description Archivistique Encodée en XML (Encoded Archival Description ou EAD)² des fiches manuscrites assistée par le logiciel de gestion d'archives, The Archivists' Toolkit (AT), ces outils et méthodes méritent d'être décrits, car c'est grâce à eux que ce corpus qui

² Pour plus de détails on pourra consulter le dictionnaire de balises EAD. C'est également le format qu'a choisi avant nous l'Institut Jean-Toussaint Desanti pour décrire les manuscrits du philosophe Jean Toussaint Desanti (L'archive numérique).

donne à voir la « bibliothèque » choisie par le philosophe pour la rédaction des *Mots et les choses* est aujourd'hui intelligible et surtout utilisable pour la recherche.

1. Les manuscrits du travail philosophique : contexte et réception de l'archive

1.1. Contexte

Dans une récente interview, Daniel Defert, compagnon du philosophe, relate la manière dont Foucault détruisait les manuscrits de ses ouvrages au fur et à mesure qu'ils étaient publiés :

J'ai cédé à la BNF le seul manuscrit qui restait, de *l'Histoire de la sexualité*. Le livre est paru quand Foucault était déjà à l'hôpital, et il n'a donc pas pu détruire le manuscrit. Sans quoi, il l'aurait détruit comme tous les autres : je l'ai vu jeter ses manuscrits, et je ne lui ai jamais dit qu'il devrait les conserver : Foucault aurait rigolé, et les aurait détruits encore plus vite.

(Bellon 4)

Parmi les documents qui échappèrent à cette destruction systématique et qui constituaient en somme une archive vivante de ses recherches, les notes venant du travail de Foucault en bibliothèque permettent non seulement de préciser l'importance de la lecture dans le processus de compréhension et de re-création d'un appareil notionnel ainsi que dans le processus d'écriture du philosophe, mais aussi certaines méthodes de travail rigoureuses. Daniel Defert d'ajouter :

Foucault avait en effet une manière de prendre des notes très spéciale : une citation par page, qu'il rangeait thématiquement. Or, corrélérer ces références au livre qu'on publie est pratiquement impossible. Les éditeurs [...] sont souvent obligés de retourner en bibliothèque, à la recherche des références de Foucault [...]. L'essentiel des archives que Foucault a laissé, ce sont ces citations, peut-être dix mille feuilles au format A4. On n'a pas forcément idée de cet énorme travail d'archive de Foucault, mais il lisait très vite, énormément : toute sa vie durant, Foucault a passé six heures par jour en bibliothèque [...].

(Bellon 5)

On sait par les différents biographes de Foucault, Didier Eribon et David Macey, que celui-ci était un des grands habitués de la rue Richelieu, mais la Bibliothèque nationale de France ne fut pas l'unique ressource du philosophe. À partir des années 1980, il se rendit également à la Bibliothèque dominicaine du Saulchoir, rue de la Glacière dans

le treizième arrondissement de Paris. Foucault a fréquenté aussi d'autres bibliothèques comme la *Carolina Rediviva* d'Uppsala ou celle de Varsovie. Dans cette liste, il en est une peut-être plus décisive, celle de sa formation, à l'École normale supérieure de la rue d'Ulm où, dès les années 1900, les élèves et professeurs pouvaient naviguer à leur guise, passant d'un ouvrage d'histoire des sciences à un volume de poésie. On ignorait également que Foucault avait porté de l'intérêt à certains ensembles de manuscrits et d'archives conservés dans d'autres bibliothèques parisiennes. L'analyse de l'un des dossiers préparatoires à *l'Histoire de la folie*, relatif à ses recherches en bibliothèque, permet d'apporter un éclairage inédit sur les pratiques informatives de Foucault, sur ses différents « savoirs-faire » ou « techniques intellectuelles » qui lui sont propres au cours de la rédaction de sa thèse. On trouve en effet dans ce dossier³ la trace de ses consultations dans différentes bibliothèques parisiennes dont la Bibliothèque historique de la ville de Paris, la Bibliothèque nationale de France (collections Clairambault et Joly de Fleury) mais aussi la Bibliothèque du duc de la Vallière, ou encore celle de Sainte Geneviève.

1.2. Réception des pièces manuscrites

Pour *Les Mots et les choses*, c'est l'ensemble du dossier préparatoire qui a été mis à disposition de notre équipe sous forme de 1712 images.tiff (soit 856 fiches numériser recto/verso), en résolution 300 DPI, taille moyenne 2550px L. * 3600px H. Ces fiches sont réparties dans cinq dossiers titrés qui, pour une large mesure, reprennent les grandes thématiques de l'ouvrage : « Analyse des richesses » (176 fiches) ; « Grammaire » (230 fiches) ; « Homme » (18 fiches) ; « Langage » (151 fiches) et « Histoire naturelle » (281 fiches). Ces fiches sont de trois types : fiches de prises de notes qui ont une disposition identique (voir Partie 2), fiches bibliographiques où le philosophe barre les titres des ouvrages vus, et fiches thématiques, plus rares, qui lui permettent de réunir sous un mot commun diverses lectures.

L'équipe de recherche en charge du dossier a entrepris, en collaboration étroite avec l'École normale supérieure de Lyon pour la partie technique,⁴ de constituer un catalogue systématique, critique et philologique⁵ de cette archive pour en conserver et en préserver

³ Ce dossier, encore en court de traitement, fait parti des notes préparatoires de Foucault. Il est relatif aux travaux bibliographique du philosophe dans plusieurs bibliothèques parisiennes, dont la Bibliothèque Nationale de France, la bibliothèque de l'Arsenal, mais aussi la bibliothèque Sainte-Geneviève et la bibliothèque historique de la ville de Paris. Un autre dossier préparatoire de *l'Histoire de la folie* est présenté sur le portail Foucault de l'IMEC.

⁴ Le projet de *La Bibliothèque Foucauldienne* a été rendu possible grâce au financement de l'Agence Nationale pour la Recherche. Il comprend les équipes suivantes : l'équipe Anthropologie de l'écriture (IIAC – EHESS – Paris), resp. Philippe Artières et l'équipe Philosophie de la politique (UMR 5206 – ENS de Lyon), resp. Jean-Claude Zancarini.

⁵ Il ne s'agit pas de chercher l'origine de chacun des travaux du philosophe, ni même de travailler sur ses brouillons, comme pourraient le faire les généticiens de la littérature mais nous intéresser à l'ensemble des

l'existence mais surtout pour étudier l'appareil notionnel et les méthodes de travail de Michel Foucault sur les sources utilisées.

Tout d'abord, grâce à la description des fiches et à leur comparaison avec les textes sources lus par Michel Foucault, il s'est agi de comprendre et d'analyser le travail existant en amont du processus d'écriture (travail bibliographique, lectures, modes de consultation des sources).

Ensuite, grâce à une comparaison à grande échelle entre ces fiches et le texte produit par Michel Foucault, il a été question de revenir sur les modes d'extractions de Foucault (comment le philosophe passe du « nomadisme » bibliographique à un travail méticuleux de commentaire de texte), sur les pratiques de restitutions des sources (quelle pratique de la bibliographie ? quelle pratique de la citation ?), et sur la construction de l'appareil notionnel et la création de concept.

Ces fiches n'ont pas fait l'objet d'une transcription systématique ni d'une réédition au format texte mais d'un inventaire détaillé, augmenté d'un index, de listes bibliographiques, de descriptions, d'explications comparatives entre les textes lus et les fiches, et entre les fiches et les textes produits. Nous allons maintenant décrire les outils et méthodes qui nous ont aidés à réaliser ce travail et à le partager.

2. Une archive numérique : outils et méthodes

Notre projet tend à un double objectif : conserver une copie numérique et intellectuelle de l'archive et créer un instrument de recherche, capable d'interroger l'ensemble de l'archive privée. Les papiers de travail de Michel Foucault ont donc fait l'objet d'une description systématique, plutôt que d'une transcription.

2.1. Pourquoi avoir décrit aussi finement l'archive ?

Le dossier physique *Les Mots et les choses* n'étant pas encore confié pour conservation à l'Institut Mémoires de l'Édition Contemporaine (IMEC), il pourrait être détruit à tout moment. Nous nous devons d'en garder une description archivistique poussée ; en nous appuyant sur sa reproduction numérique. Or l'archivistique, a, depuis les années 1960, bénéficié, au même titre que la bibliothéconomie et la philologie, des évolutions du numérique. En effet, la création des modèles structurés d'échange et de traitement bibliographique, les formats Machine Readable Cataloging (MARC) en 1965, Universal Machine Readable Cataloging (UNIMARC ; Willer) en 1977, le standard Text Encoding Initiative (TEI ; Burnard) en 1987 qui, lui, propose des guidelines de méthode d'encodage pour la mise en ligne de texte, et enfin le format EAD en 1993, inspiré de la TEI et des

traces qui témoignent de ses pratiques de lecture, d'enquêter sur ses prises de note en bibliothèque, sur ses manières d'archiver, mais aussi sur la fonction citationnelle, les notes de bas de page (Grafton), l'usage de la bibliographie, et l'ensemble du paratexte (Genette) – sans négliger l'examen de sa propre bibliothèque.

formats MARC, sont autant d'évolutions qui permettent aujourd'hui de répondre aux besoins de notre projet.

2.2. Encodage : EAD vs TEI msDesc

Arrêtons-nous ici sur les raisons qui nous ont fait choisir le standard adapté à l'archivistique, EAD, plutôt que le standard classique de la philologie numérique⁶, le standard TEI⁷.

En effet, le projet Européen mené par Peter Robinson entre 1999 et 2001 *Manuscript Access through Standards for Electronic Records* (MASTER) offre lui aussi un modèle TEI de description de sources primaires manuscrites très intéressant : *TEI Manuscript Description* ou *TEI msDesc*. Pour choisir entre TEI msDesc et EAD, nous avons pris en compte trois critères.

Un encodage pour tous les niveaux de l'archive

Tout d'abord le modèle msDesc, s'il prévoit un niveau de description très fin d'une source primaire, ne rend pas bien compte de l'ensemble dans lequel la source existe. En effet, le niveau de description le plus haut prévu par ce modèle est l'élément <msDesc/>, ce qui correspondrait à notre niveau le plus bas, <c level="item"> en EAD. Ce modèle prévoit néanmoins la possibilité de décrire une collection de pièces manuscrites, en intégrant le niveau <msDesc> dans l'élément bibliographique <listBibl> : « However, in cases where the document being encoded is essentially a collection of manuscript descriptions, the msDesc element may be used in the same way as the bibliographic elements (bibl, biblFull, and biblStruct) making up the TEI element class model.biblLike. These typically appear within the listBibl element » (10.2).

Avec ce modèle, la description de l'archive serait placée dans l'en-tête du fichier, avec une transcription ou des images dans le corps du fichier XML :

```
<TEI.2>
  <teiHeader>
    <fileDesc>
      <sourceDesc>
        <msDesc/>
      </sourceDesc>
```

⁶ En nous appuyant sur les définitions de la philologie numérique de Franck Neveu (226–27) et François Rastier (73), nous entendons ici, par philologie numérique, l'étude informatisée des textes manuscrits qui nous permettront de découvrir les processus et procédés d'écriture du philosophe (ici, ceux de Michel Foucault) : des lectures à l'appropriation de champs disciplinaires et notionnels ; de l'appropriation à la rédaction, comment les manuscrits laissés par le philosophe nous permettent de comprendre comment il écrivait. Cette étude est dite informatisée à deux titres : premièrement les manuscrits et leur description sont annotés/étiquetés pour permettre une automatisation des recherches. D'autre part, l'application finale qui permettra la visualisation des résultats devra être modélisée afin d'être développée et/ou optimisée.

⁷ Initiative à l'origine de la création de guides de méthodes d'encodage de textes pour les sciences humaines et sociales.

```

</fileDesc>
</teiHeader>
<text>
  <pb facs="image1" />
</text>
</TEI.2>

```

Or, si nous considérons notre archive, c'est bien l'ensemble des pièces qui doit être décrit et interrogeable dans son organisation originale. On doit pouvoir interroger n'importe quel niveau de l'ensemble, soit au moins six niveaux : le fonds, les groupes de documents, les séries, les dossiers, les sous-dossiers, et les pièces. En adaptant les éléments <listBibl/>, corpus et <msDesc/>, nous y serions parvenu, mais nous avons, pour des raisons de temps impartis très courts, choisi d'utiliser le format EAD qui offrait un jeu d'éléments et d'attributs adaptés à l'organisation et la nature sérielle de notre archive. Nous reviendrons sur le détail des balises EAD choisies pour notre description, dans la partie 2.3.

Des outils plus pratiques

Voyons maintenant notre deuxième critère : EAD nous offre une palette d'outils facilement utilisables et paramétrables que n'offre pas encore la TEI. Les logiciels de gestion d'archives permettent de renseigner et décrire l'archive dans des interfaces plus accessibles aux chercheurs en sciences humaines et sociales que des éditeurs de code : Digitool, ICA-Atom, The Archivists' Toolkit (AT), Archon (Spiro) etc. Certains de ces logiciels offrent également la possibilité de publier sur le web l'archive interrogeable (index de personnes, index thématique, recherche plein texte). Suite à des tests d'exports XML très propres et satisfaisants, nous avons donc finalement opté pour l'Archivists' Toolkit. C'est une application JAVA Client/Serveur élaborée par l'*University of California San Diego Libraries*, la *New York University Libraries* et la *Five Colleges, Inc. Libraries*. Ce projet a été financé par l'*Andrew W. Mellon Foundation*. Régulièrement, l'archive est exportée au format EAD XML pour être corrigée ou adaptée par les ingénieurs qui utilisent l'éditeur de code oXygen.

Correspondance avec les formats bibliographiques et les normes archivistiques

Les logiciels de gestion d'archives assurent la correspondance entre EAD et le format MARC ; et entre EAD et la Norme Générale Internationale de description archivistique ISAD(G). En effet les correspondances entre ces différents formats font partie du processus de création et de mise à jour de la DTD 2002 de l'EAD.

Pour ces 3 raisons notre choix s'est donc porté sur EAD. La TEI sera utilisée, au cas par cas, quand un membre de l'équipe décidera de transcrire telle ou telle pièce : les transcriptions seront enregistrées et stockées comme des objets numériques (au même titre que des images) : une simple déclaration et pointeur dans l'élément <dao/> du fichier TEI suffiront, comme nous le verrons plus loin.

Ainsi, le format et les outils choisis nous ont donné la possibilité de décrire l'archive pièce par pièce, de rendre compte des regroupements (thématiques, typologiques, etc.) effectués par Foucault et de les déplier, de rendre compte du découpage et de l'organisation interne d'une fiche, de créer des index et de repérer dans un document des termes ou notions clefs, c'est à dire de donner à voir le contenu notionnel des sources d'après la lecture qu'en a fait Foucault.

2.3. EAD et le choix des niveaux de description

En respectant le dictionnaire de balises EAD, la DTD 2002 et le manuel d'encodage publiés respectivement en 2002 et 2005 par le groupe CG46/CN357/GE3⁸ de l'Agence Française de Normalisation (AFNOR), nous avons choisi les niveaux de description qui pouvaient le mieux représenter l'archive existante, tout en laissant ouverte la possibilité de tout type d'ajouts ultérieurs et d'ajouter des niveaux de description ultérieurs.

En voici une description accompagnées d'illustration. Nous retenons donc pour l'instant les niveaux suivants :

- Fonds : Les papiers de travail de Michel Foucault
- Recordgp ou Sous-fonds : *Les mots et les choses*
- Series : Les fiches de lecture
- File : Analyse des richesses, etc.

Le fonds <archdesc level="fonds">

Les papiers de travail de Foucault (séparés entre plusieurs lieux dont la BNF, l'appartement de Daniel Defert, celui d'Arlette Farge, l'IMEC, etc.) pourraient faire considérer l'ensemble comme une archive artificielle et *non-organique*. Or, c'est le processus de production des pièces qu'il faut considérer, et non leur localisation, pour juger si l'ensemble est *organique* ou *non-organique*.

En EAD, on considère comme *organique* un ensemble de pièces homogènes au moment de la production, par exemple un ensemble de pièces produites par un même auteur, ou une même institution. Les pièces à décrire ici, « *les papiers de travail de Michel Foucault* », forment un ensemble organique au moment de leur production, puisqu'ils proviennent du même auteur : nous avons donc choisi comme niveau de description le plus haut : le *fonds*. Une *collection* en revanche, serait, par exemple, un ensemble de pièces disparates (différents types de pièces ou des pièces produites par différents producteurs, etc.) réunies artificiellement autour d'un sujet.

⁸ Parmi ce groupe de travail AFNOR, Denise Ogilvie (Archives Nationales de France) nous a accordé un entretien très utile à ce sujet.

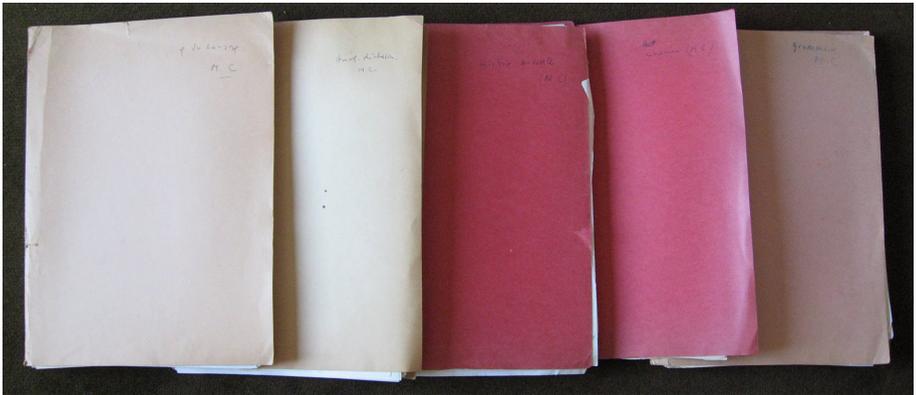


FIGURE 1. Ces 5 dossiers correspondent aux grands chapitres de *Les Mots et les Choses* : « Philosophie du langage », « Analyse des Richesses », « Histoire Naturelle », « Homme », « Grammaire ».

Le groupe de documents <c level="recordgrp">

Les Mots et les choses représente un groupe de documents appartenant au fonds Papiers de Michel Foucault. Il contient tous types de papiers de travail ayant été utilisés par Michel Foucault pour rédiger *Les Mots et les choses*. La raison pour laquelle le niveau de description *Les mots et les choses* a été choisi comme englobant les fiches de lecture sera détaillée plus bas, avec la description des « dossiers ».

Les séries <c level="series">

En EAD, le terme *series* s'utilise pour des pièces qui sont produites à la chaîne (par exemple un formulaire type produit en série par une institution). Il apparaît donc logique d'utiliser *series* pour désigner l'ensemble des fiches de lecture et non plus pour désigner l'ensemble *Les mots et les choses*.

Les dossiers <c level="file">

Nous avons choisi *Les mots et les choses* comme un groupe de documents englobant les fiches de lecture contenues dans les 5 dossiers (« Philosophie du langage », « Analyse des Richesses », « Histoire Naturelle », « Homme », « Grammaire ») (cf. fig. 1) qui correspondent aux thématiques annoncées dans la préface de l'ouvrage : « pour former le socle positif des connaissances telles qu'elles se déploient dans la *grammaire* et dans la *philologie*, dans l'*histoire naturelle* et dans la biologie, dans l'*étude des richesses* et dans l'économie politique. [...] Mais à mesure que les choses s'enroulent sur elles-mêmes, ne demandant qu'à leur devenir le principe de leur intelligibilité et abandonnant l'espace

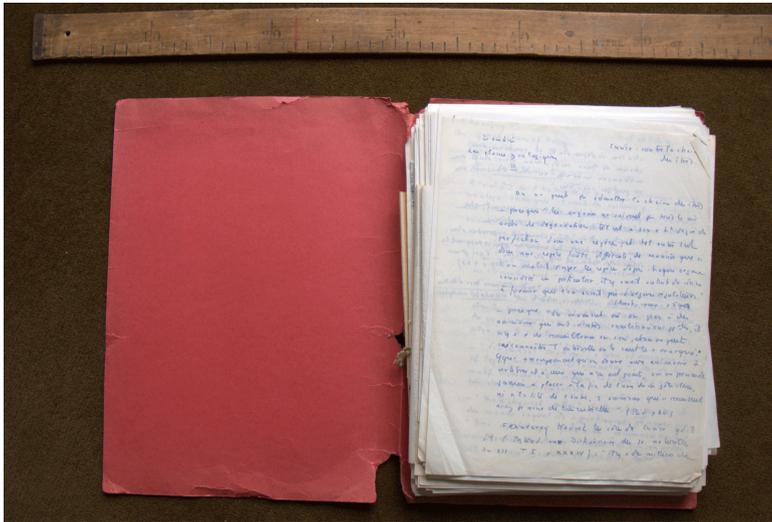


FIGURE 2. Le dossier « Histoire Naturelle ».

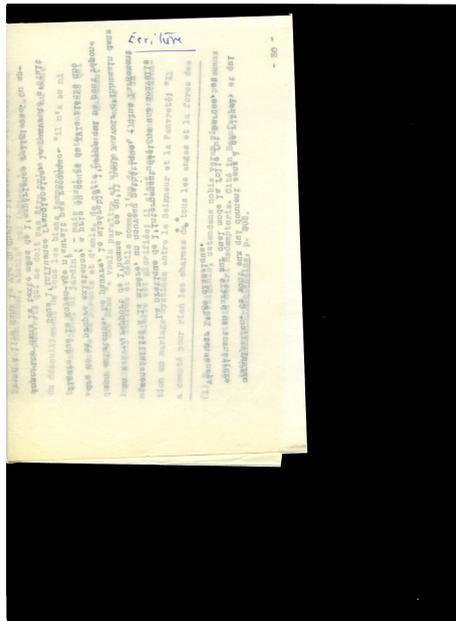


FIGURE 3. Page A4 pliée en deux : un sous-dossier présent dans le dossier « Grammaire » : « Écriture ».

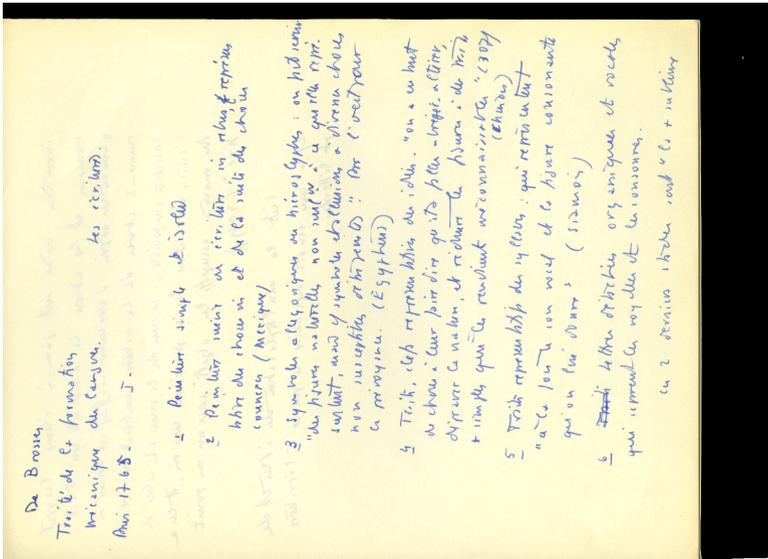


FIGURE 4. Une fiche de lecture du sous-dossier « Écriture ».

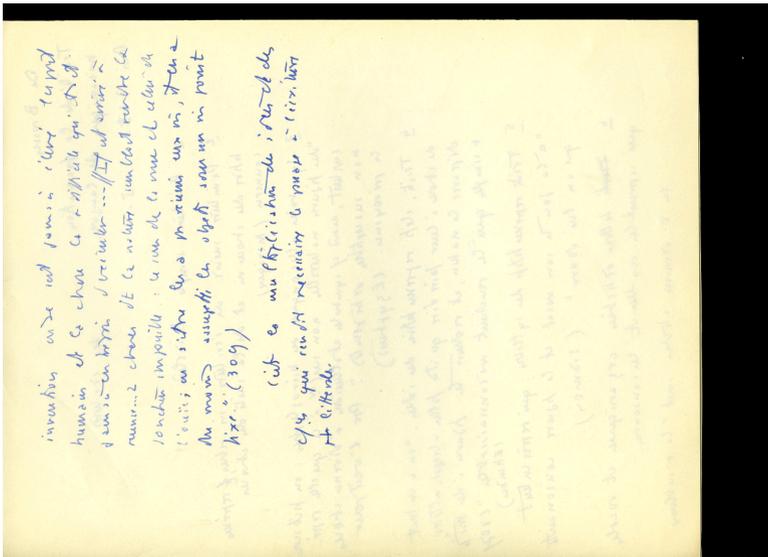


FIGURE 5. Une fiche de lecture du sous-dossier « Écriture ».

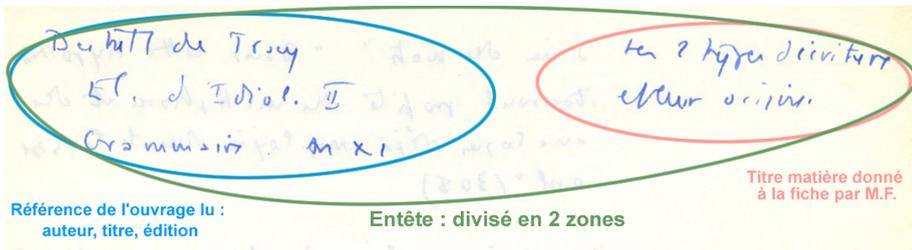


FIGURE 6. En-tête d'une fiche de lecture (1).

de la représentation, l'*homme* à son tour entre, et pour la première fois, dans le champ du savoir occidental » (13–15).

En consultant le sommaire, le rapprochement est encore plus net et permet de déduire les correspondances suivantes :

Le chapitre un, *La prose du monde* correspond au dossier « Grammaire » ; le chapitre deux, *Parler* correspond au dossier « Philosophie du langage » ; le chapitre trois, *Classer* correspond au dossier « Histoire naturelle » ; le chapitre quatre *Échanger* correspond au dossier « Étude des richesses /économie » ; le chapitre cinq, *L'homme et ses doubles* correspond au dossier « Homme ».

Ces sous-ensembles sont considérés comme des *dossiers* (<c level="file">) qui sont englobés dans une *série* (<c level="series">) de *fiches de lectures*, qui sont elles-mêmes englobées dans un *ensemble de documents* (<c level="recordgrp">) ayant servi à la rédaction de l'ouvrage *Les Mots et les choses*.

Les sous-dossiers <c level="file">

Les dossiers sont divisés en plusieurs sous-dossiers, que nous considérerons comme des *dossiers* (<c level="file">).

La pièce = une fiche de lecture <c level="item">

Le niveau de description le plus bas (celui qui va demander le plus de minutie et d'attention), est celui de la pièce : la fiche de lecture.

Foucault consacre une page recto-verso pour citer un passage d'un ouvrage qu'il cite sous un titre matière. Cette pratique permet à Michel Foucault de découper l'ouvrage à partir de ses propres problématiques. Pour chaque fiche, Foucault nous donne au moins :

- le nom de l'auteur ;
- le titre de l'ouvrage et l'année de publication (souvent la première édition) ;
- un titre matière qui lui permet de regrouper les passages cités.

En voici un premier exemple sur la figure 6. En voici un deuxième, grâce à la figure 7.

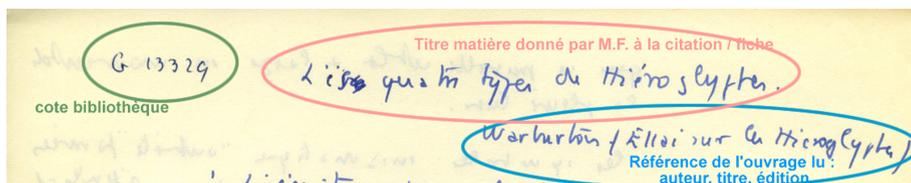


FIGURE 7. En-tête d'une fiche de lecture (2).

La version 2002 de la DTD EAD nous offre en effet la possibilité de renseigner trois niveaux de description des fiches : un niveau général regroupant le titre de la fiche, la langue, le numéro d'inventaire des documents quand celui-ci est renseigné par Foucault ; un niveau biographique et thématique qui permet de renseigner le nom de l'auteur et les différents thèmes abordés par Foucault dans chacune des fiches – il a fallu pour cela élaborer un vocabulaire thématique à partir des fiches ; enfin un niveau bibliographique et analytique qui permet de renseigner la référence de l'ouvrage cité, et d'analyser le document à partir de plusieurs points d'entrée (mots soulignés, expression traduite par Foucault, pagination des citations). Chaque fiche a aussi fait l'objet d'une description matérielle avancée dans laquelle sont renseignées la nature du support et la graphie. Voici comment nous récupérerons ces diverses informations :

par la balise de titre : `<unittitle></unittitle>` peut elle-même contenir plusieurs balises de titre. On peut donc utiliser pour décrire un des deux cas de figure ci-dessus : `<unittitle>Destutt de Tracy. Element d'Idéologie II. Grammaire - XI. <title type="TM_MF">Les deux types d'écriture et leur origine</title></unittitle>`.

par l'indexation : pour créer les index de « mots sujets », ou « titres matières », plutôt qu'utiliser le langage d'indexation Rameau, nous avons trouvé plus intéressant d'utiliser le vocabulaire de Michel Foucault pour donner à voir la construction d'un appareil notionnel : cela nous permettra d'indexer le contenu propre au corpus. A cet effet, on indexera, dans les fiches, les titres matières (à gauche souvent, dans la zone de titre), et les mots mis en évidence par Michel Foucault.

2.4. Fichier EAD et objets numériques

Il est intéressant d'observer dans l'encodage EAD XML (fig. 8) d'une pièce comment les objets (livres de référence lus par Michel Foucault, fiches de lecture, ouvrages produits par Michel Foucault, images obtenues par la numérisation) sont liés au fichier de

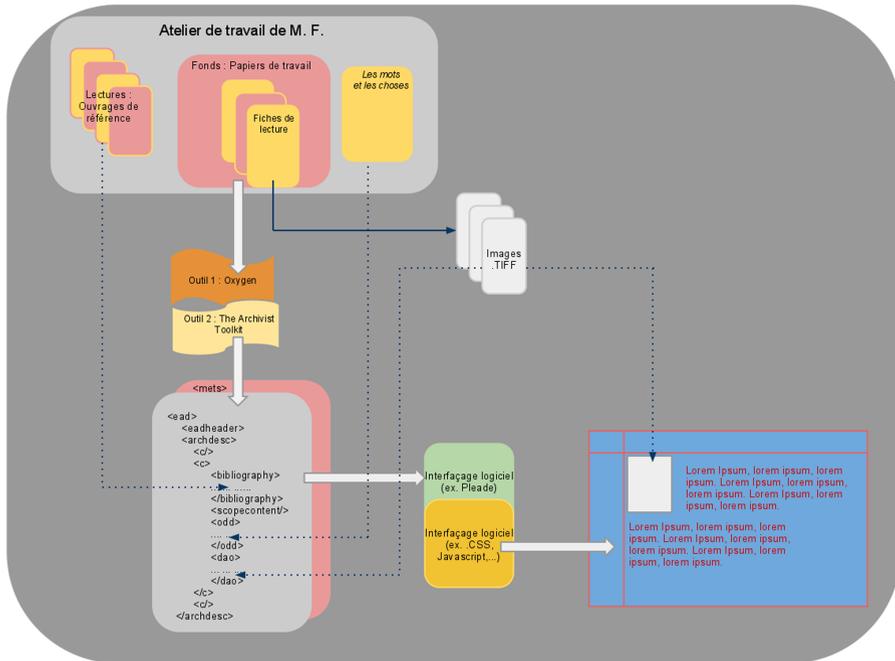


FIGURE 8. Un exemple simplifié de la description d'une pièce (<c level="item">).

description EAD XML.

Dans la balise <bibliography/> ce sont toutes les références nécessaires pour retrouver la littérature à l'origine des fiches que l'encodeur spécifiera.

Dans la balise <odds> ce sont les ouvrages et les passages produits par Michel Foucault grâce à la fiche de lecture qui seront spécifiés. Enfin, les liens vers les différentes versions de la fiche, c'est-à-dire ici vers les images numérisées correspondantes (mais il pourrait tout aussi bien s'agir d'une transcription TEI), seront faits dans le balise <dao/> : l'objet archivistique numérique. Nous envisageons d'utiliser cette balise pour les fiches que les chercheurs décideront de transcrire en TEI.

2.5. L'application finale pour présenter les résultats : modélisation UML

Les outils utilisés jusqu'ici ne suffisent pas à rendre ergonomique, interactive et hautement intuitive l'application finale, ni à définir pour cette même application, les

```

<c id="ref43" level="item">.
  <did>.
    <umittitle>Destutt de Tracy. Element d'Idéologie II. Grammaire -- XI. .
      <title type="TM_MF">Les deux types d'écriture et leur origine</title></umittitle>.
      <unitid>0007</unitid>. TM_MF = Titre matière donné par Michel Foucault
      <langmaterial>.
      <language langcode="fre"/>.
      </langmaterial>.
      <physdesc>.
      <extent>2.0 digitized images</extent>.
      </physdesc>.
    </did>.
    <bibliography id="ref45">. spécification de la référence de l'ouvrage lu faisant l'objet de la fiche
      <head>Référence de l'ouvrage</head>.
      <bibref>Destutt de Tracy, Eléments d'idéologie. Seconde partie. Grammaire, Paris : Courcier, an XI-1803.
    </bibliography>.
    <scopecontent id="ref46">.
      <head>Analyse du document</head>.
      <p>Mot(s) souligné(s) :</p>.
      <p>Michel Foucault nous donne ici la cote de l'ouvrage emprunté pour la réalisation de cette fiche :</p>.
      <p>Les passages cités par Foucault dans cette fiche sont issues des pages : p. 307, p. 308, p.309 et p. 3
    </scopecontent>.
    <odd id="ref47">.
      <head>Commentaires</head>. ouvrage et passages produits par Foucault utilisant cette référence / citation
      <p>Traduit par Foucault de la langue originale :</p>.
      <p>Michel Foucault cite, dans <i>Les mots et les choses</i>, l'auteur faisant l'objet de cette fiche : le
      <p>Cet auteur est cité dans cet autre ouvrage de Michel Foucault :</p>.
    </odd>.
    <dao ns2:actuate="onRequest" ns2:show="embed" ns2:title="page 01 - recto" ns2:href="PMF_MC_FL_Gr_Ec_HR_003">.
      <daodesc>.
      <p>page 01 - recto (PMF_MC_FL_Gr_Ec_HR_003)</p>. lien vers un objet, ici un Digital Object :
      </daodesc>. une image de la fiche manuscrite.
    </dao>. il pourrait s'agir d'une transcription (ex. 1 fichier TEI)
    <controlaccess>.
      <persname source="ingest">Destutt de Tracy, Antoine-Louis-Claude (1754-1836)</persname>
      <subject source="VNF">hiéroglyphes</subject>.
      <subject source="VNF">origine de l'écriture</subject>. les index persname et subject
      <subject source="VNF">types d'écriture</subject>.
      <subject source="VNF">égyptien ancien (langue) -- écriture hiéroglyphique</subject>.
    </controlaccess>.
  </c>.

```

FIGURE 9. Diagramme de navigation pour l'internaute invité.

différents niveaux d'accès aux reproductions iconographiques, dans le respect de la volonté des légataires. Nous venons donc d'entrer dans une nouvelle phase du projet, celle de la modélisation UML de l'application web finale. Cette modélisation s'appuie sur la démarche définie par Pascal Roques dans le premier chapitre de son ouvrage (2-21). Elle nous permettra de définir les acteurs, les cas d'utilisation, la maquette, ainsi que tous les diagrammes (les diagrammes de séquences système, de classes participantes, des interactions, de navigation, et de classes de conception) nécessaires au choix d'un logiciel (Pleade)⁹, ou au développement d'une application.

Ces diagrammes nous permettront d'envisager l'application finale, jusqu'au détail le plus fin. On verra à titre d'exemple le diagramme de navigation d'un internaute

⁹ Exemples de logiciels envisagés pour la mise en ligne : Pleade (AJLSM 2009).

« invité », c'est à dire de l'internaute bénéficiant d'identifiants de connexion pour une navigation simple avec un accès aux textes et aux images (cf. fig. 9). Il nous faudra définir ainsi les diagrammes nécessaires pour chaque type d'utilisateur, et les synthétiser jusqu'à obtenir un modèle de conception détaillé.

Conclusion

Les apports du numérique ont été essentiels à ce travail de recherche sur les papiers de travail de Michel Foucault. En effet, ni l'annotation assistée de ses fiches de lecture manuscrites, ni l'automatisation des recherches dans l'archive et les objets qui lui sont liés (références d'ouvrages, ouvrages produits, images numérisées), ni la modélisation de l'application finale qui offrira un pan de la bibliothèque foucauldienne aux yeux des internautes, n'auraient été possibles sans les apports méthodiques et logiciels de l'informatique à la philologie du XXI^{ème} siècle. Pourtant, ces « apports » ont également été sources de questionnements difficiles, comme le choix du format EAD plutôt que le standard TEI. Pour cette question précise, c'est, comme nous l'avons vu plus haut, EAD, une norme compatible aux formats bibliographiques, aux normes archivistiques et permettant une exploitation logicielle plus adaptée et simplifiée qui l'a emporté.

Dans l'état actuel d'avancement du travail de description, le traitement archivistique de ces fiches nous a permis d'envisager de nouvelles hypothèses de travail autour du travail foucauldien et tout d'abord de reprendre la question du renversement des hiérarchies et le choix délibéré de Foucault de choisir certains auteurs dits « secondaires » pour appréhender des champs de savoir parfois nouveaux pour lui (exemple : économie).

Nous avons également pu mieux comprendre le degré des citations. En effet, Foucault cite peu par rapport à ce qu'il a lu et surtout noté. Les exemples sont nombreux d'auteurs cités dans plusieurs fiches mais qui n'apparaissent pas dans l'ouvrage publié (c'est le cas par exemple du révérend père Coeurdoux connu pour son analyse de la ressemblance entre le sanscrit et le latin. Si son nom apparaît dans 4 fiches, Foucault ne le cite qu'en page 305, dans le corps du texte sans en dire plus).

Au-delà de ces premières hypothèses concernant le travail de Foucault, d'autres hypothèses sont envisageables concernant cette fois-ci le travail intellectuel en général. Foucault est pour cela un « laboratoire » de première importance. Ce type de travail sur les archives de la recherche permettrait par exemple de proposer une approche inédite du travail philosophique au XX^e siècle qui rendrait compte des principes de transformations de l'attitude rhétorique des intellectuels dans leur manière de présenter leurs idées, de les mettre en œuvre et de les diffuser : à l'université, devant un public cultivé, ou à la radio, à la télévision. Ce travail sur les archives permet aussi de mieux comprendre la phase de professionnalisation de la philosophie française après la seconde

guerre mondiale, qui se définit entre autre par une porosité plus grande vers d'autres disciplines dont l'histoire et les sciences sociales.

Enfin, l'on peut également à partir du cas Foucault reprendre la question de la créativité intellectuelle et de la transmission d'un style, ou de la manière de traiter les idées. Réflexion importante puisqu'il semble, encore, que l'originalité d'un produit textuel « philosophique » soit liée, en grande partie, à la méconnaissance de son processus de fabrication.

Bibliographie

- L'archive numérique Jean-Toussaint Desanti*. Lyon : ENS de Lyon, UMR 5037 Institut d'Histoire de la Pensée Classique, 2010. <<http://institutdesanti.ens-lyon.fr/spip.php?rubrique4>>.
- AT : *Archivists' Toolkit*. 2006–2009. <<http://archiviststoolkit.org/>>.
- Bellon, Guillaume. « Je crois au temps... » Daniel Defert, légataire des manuscrits de Michel Foucault, propos recueillis par Guillaume Bellon. » *Recto/Verso* 1 (2007) : 1–7. <<http://www.revuerectoverso.com/spip.php?article29>>.
- Bellour, Raymond. *Le livre des autres, essais et entretiens*. Paris : L'Herne, 1971.
- Bustarret, Claire. « Approche codicologique du manuscrit moderne. » *Critiques génétique, concepts, méthodes, outils*. Ed. Olga Anokhina et Sabine Petillon. Paris : IMEC Editeur, 2009. 47–60.
- Burnard, Lou. *Text Encoding for Interchange : A New Consortium*. 2000. <<http://www.ariadne.ac.uk/issue24/tei>>.
- EAD : *Description Archivistique Encodée. Dictionnaire de balises*. Trad. de l'anglais par le groupe AFNOR CG46/CN357/GE3. Society of American Archivists. (SAA), 2004. <<http://www.archivesdefrance.culture.gouv.fr/static/1066>>.
- Foucault, Michel. *Dits et écrits, T. IV*. Ed. Daniel Defert et François Ewald. Paris : Gallimard, 1995.
- Foucault, Michel, *Histoire de la folie, Extraits du dossier manuscrit préparatoire*. Portail Michel Foucault. Caen : IMEC, 2010. <<http://michel-foucault-archives.org/spip.php?article17>>.
- Foucault, Michel. *Les mots et les choses : une archéologie des sciences humaines*. Paris : Gallimard, 1990.
- Genette, Gérard. *Palimpsestes : La littérature au second degré*, Paris : Éditions du Seuil, 1982.
- Grafton, Anthony. *Les origines tragiques de l'érudition. Une histoire de la note en bas de page*. Paris : Éditions du Seuil, 1998.
- IMEC : *Institut Mémoires de l'Édition Contemporaine*, 2010. <<http://www.imec-archives.com/>>.
- ISAD(G) : *Norme générale et internationale de description archivistique : adoptée par le Comité sur les normes de description, Stockholm, Suède, 19–22 septembre 1999*. Ottawa : International council on archives (ICA), 2000. <<http://www.ica.org/fr/node/30001>>.
- Jacob, Christian. *Lieux de savoir, Vol. 1. Espaces et communautés*. Paris : Editions Albin Michel, 2007.
- Kraus, Dorothea. « Appropriation et pratiques de la lecture. Les fondements méthodologiques et théoriques de l'approche de l'histoire culturelle de Roger Chartier. » *Labyrinthe* 3 (1999). 13–25. <<http://labyrinthe.revues.org/index56.html>>.

- Latour, Bruno. « Les vues de l'esprit, une introduction à l'anthropologie des sciences et des techniques. » *Culture technique* (1985) : 4–30.
- Manuel d'encodage : Faire un répertoire ou un inventaire simple en EAD (Description archivistique encodée)* : Groupe d'experts AFNOR CG46/CN357/GE3. 2005–2009. <<http://www.archivesdefrance.culture.gouv.fr/static/3322>>.
- MASTER : *Manuscript Access through Standards for Electronic Records. Reference Manual for the MASTER Document Type Definition – Discussion Draft*. Ed. Lou Burnard for the MASTER work Group. Oxford : Oxford University Computing Services, 2001. <http://www.tei-c.org/About/Archive_new/Master/Reference/oldindex.html>.
- Müller, Bertrand. « Les lieux de savoir. Un entretien avec Christian Jacob. » *Genèses* 76 (2009) : 116–136.
- Neveu, Franck. *Dictionnaire des sciences du langage*. Paris : Armand Colin, 2004.
- Pleade*. Bordeaux : Societé AJSLM, 2009–2010. <<http://pleade.com/>>.
- Rastier, François. *Arts et sciences du texte*. Paris : Presses Universitaires de France, 2001.
- Roques, Pascal. *UML. Modéliser une application web*. Paris : Eyrolles, 2008.
- Spiro, Lisa. *Archival Management Software. A Report for the Council on Library and Information Resources*. January 2009. <<http://www.clir.org/pubs/reports/spiro2009.html>>.
- TEI : *Text Encoding Initiative*. TEI Consortium, 2010. <<http://www.tei-c.org>>.
- Manuscript Description* : <<http://www.tei-c.org/release/doc/tei-p5-doc/fr/html/MS.html>>.
- Willer, Mirna. *UNIMARC Manual*. München : K. G. Saur, 2007.