



Genesis

Manuscrits – Recherche – Invention

49 | 2019

Une tradition italienne

Le « siliscrit » de Jacques Derrida. Exploration d'une archive nativement numérique

Aurèle Crasson, Jean-Louis Lebrave et Jérémy Pedrazzi



Édition électronique

URL : <https://journals.openedition.org/genesis/4316>

DOI : [10.4000/genesis.4316](https://doi.org/10.4000/genesis.4316)

ISSN : 2268-1590

Éditeur :

Presses universitaires de Paris Sorbonne (PUPS), Société internationale de génétique artistique littéraire et scientifique (SIGALES)

Édition imprimée

Date de publication : 15 décembre 2019

Pagination : 7-12

ISBN : 979-10-231-0650-3

ISSN : 1167-5101

Référence électronique

Aurèle Crasson, Jean-Louis Lebrave et Jérémy Pedrazzi, « Le « siliscrit » de Jacques Derrida. Exploration d'une archive nativement numérique », *Genesis* [En ligne], 49 | 2019, mis en ligne le 15 décembre 2019, consulté le 10 juin 2021. URL : <http://journals.openedition.org/genesis/4316> ; DOI : <https://doi.org/10.4000/genesis.4316>

Tous droits réservés

Le «siliscrit» de Jacques Derrida. Exploration d'une archive nativement numérique*

Aurèle Crasson, Jean-Louis Lebrave, Jérémy Pedrazzi

*Sous chaque feuille s'ouvrent les lèvres d'une blessure,
pour laisser entrevoir la possibilité abyssale d'une autre profondeur
promise à l'excavation archéologique**.*

Jacques DERRIDA

La préhistoire

Le projet Derrida hexadécimal n'aurait sans doute pas vu le jour sans une conférence tenue par Lou Burnard en janvier 2011 dans le cadre du séminaire général de l'ITEM. Invité pour parler de la TEI (*Text Encoding Initiative*) dont il est un des grands prosélytes, Lou Burnard avait choisi de parler d'un autre thème, la *computer forensics*, qui lui paraissait de première importance pour la critique génétique, et ce en présentant un rapport commandé par l'Office fédéral des bibliothèques des États-Unis et diffusé en décembre 2010 sur l'exploitation des fonds d'archives nativement numériques déposés dans les bibliothèques américaines¹. À la lecture de ce texte, Jean-Louis Lebrave a associé celle du livre *Mechanisms* de Matthew Kirschenbaum, coauteur du rapport, et celle d'un article publié quelques mois plus tôt par Thorsten Ries dans la revue des éditeurs critiques allemands *Editio*². Il en est résulté une présentation critique de ces trois textes dans *Genesis* n° 33³.

Le thème des *computer forensics* appliquées à des disques durs d'écrivains a fait résurgence à la faveur d'une conférence donnée par Thorsten Ries à l'École normale supérieure en 2012⁴, mais le terrain n'était visiblement pas encore mûr en France pour une mise en application sur un corpus «réel». Il a fallu attendre la préparation du Congrès du cinquantenaire de l'ITEM⁵ pour que la proposition de consacrer quelques heures à cette «codicologie des disques durs» rencontre la possibilité, grâce à Aurèle Crasson, de disposer du fonds numérique de Jacques Derrida déposé à l'IMEC comme corpus d'étude. Ce qui pourrait passer pour

un choix par défaut du fait de la rareté de ce type de dépôt dans les centres d'archives français est apparu d'emblée comme une chance extraordinaire, puisqu'on était à même ainsi d'explorer les ressources d'un auteur pour qui la question de l'archive a toujours été centrale. On quitte dès lors la préhistoire du projet pour entrer dans sa mise en œuvre. Il s'agit de développer une branche nouvelle de la codicologie à partir de l'exploitation des traces numériques (fichiers dont la datation est possible, fichiers effacés dont l'espace qu'ils occupent sur le disque n'a pas été recouvert, texte effacé dans un fichier, mais conservé en tout ou partie

(*) Nous remercions chaleureusement Laurent Alonso pour son aide dans l'interprétation des données numériques, puis pour la relecture attentive du texte et son commentaire.

(**) Jacques Derrida, *Mal d'archives*, Paris, Galilée, 1995, p. 39.

1. Matthew Kirschenbaum, Richard Ovenden et Gabriela Redwine, *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*, Washington, D.C., Council on Library and Information Resources, décembre 2010.

2. Matthew Kirschenbaum, *Mechanisms. New Media and the Forensic Imagination*. Cambridge, Mass. et Londres, MIT Press, 2008. Thorsten Ries, «"die geräte klüger als ihre besitzer". Philologische Durchblicke hinter die Schreibszenen des Graphical User Interface», *Editio*, n° 24, 2010.

3. Jean-Louis Lebrave, «Computer forensics : la critique génétique et l'écriture numérique», *Genesis*, n° 33, 2011, p. 137-147. Disponible à l'adresse suivante : <https://journals.openedition.org/genesis/633>.

4. Thorsten Ries, «Digital Archives and Genetic Textual Criticism. Perspectives of Digital Forensics for Textual Scholarship», conférence invitée, ITEM, 10 mai 2012.

5. Il s'est tenu du 17 au 20 octobre 2018 à l'École normale supérieure et à la Bibliothèque nationale de France (actes à paraître). <http://item-50ans.org/>.

pour des raisons liées à l'état de la technologie informatique de l'époque) dans les disques durs de Jacques Derrida, et d'analyser la logique derridienne de l'archive confrontée au support numérique en la comparant à ses analogues manuscrits et tapuscrits.

Le fonds

Les archives numériques de Jacques Derrida sont conservées dans les collections de l'IMEC, où elles ont rejoint son fonds papier. Les plus anciennes remontent à 1985, année de l'acquisition de son premier ordinateur et leur contenu est de nature diverse : des séminaires donnés à l'EHESS de 1985 à 2003, des cours et conférences prononcés dans le monde entier, ainsi que les œuvres publiées par Derrida depuis cette fin des années 1980 ; à cela s'ajoutent des correspondances institutionnelles et administratives, ainsi que des fichiers de listes d'adresses tant personnelles que professionnelles, amis, éditeurs, universitaires, etc., mais on n'y trouve ni images, ni courriers électroniques. Suite à un malheureux cambriolage, Derrida, hanté par la crainte de perdre le fruit de son travail, a multiplié les copies de toutes ses productions ; les archives attestent de ce foisonnement documentaire par la présence de fichiers portant le même nom et dont les dates supposées de création et les empreintes numériques⁶ permettent de penser qu'il s'agit de sauvegardes.

Depuis 2004, année du décès de l'auteur, les versements d'archives sont effectués par son épouse, Marguerite Derrida.

Il est important de souligner que si la bibliothèque de l'Université d'Irvine et l'IMEC ont accueilli notamment des archives *issues de documents numériques* (tirages sur papier de fichiers texte) du vivant de l'auteur, Jacques Derrida n'avait, en revanche, jamais de lui-même, confié d'archives *nativement numériques* (des contenus dépendant d'une structure logique et d'une machine capable de les interpréter). C'est dans le cadre d'un enrichissement du fonds effectué par Marguerite Derrida en 2008 que l'IMEC collecte au domicile de Jacques Derrida trois ordinateurs : un Macintosh Plus, un Macintosh Classic et un Macintosh LC475, ainsi que des supports externes de mémoire. Dès 2010, l'IMEC procède à une copie complète des données. Il sera plus tard signalé dans un document interne « l'importance de documents numériques traçant une première perspective de recherche à partir des fichiers sauvegardés ».

D'autres versements seront effectués par les ayants droit, notamment un volume de quelque 500 disquettes. Mentionnons au passage que le fonds est très certainement incomplet : Derrida évoque dès *Mal d'archives* (1995) un « petit Macintosh portable », et il mentionne dans une des versions de la 14^e séance du séminaire qu'il tient à l'EHESS sur la peine de mort entre 1999 et 2001 un IMAC qui contiendrait la bonne version du texte ; comme il n'a cessé d'écrire jusqu'à sa mort, il est fort à parier qu'il ait bénéficié d'un ordinateur plus récent que ceux déposés à l'IMEC.

En 2017, une première convention de partenariat est signée entre les ayants droit de Jacques Derrida et l'École normale supérieure dans le cadre d'un projet Labex initié par Aurèle Crasson, « Le Macintosh Plus de Jacques Derrida, exploration génétique d'un disque dur », en prenant acte de la volonté des parties de collaborer pour donner accès aux archives nativement numériques du philosophe. L'objet de la convention était d'appliquer les méthodes de la *computer forensics*⁷ aux archives numériques de Jacques Derrida. Aurèle Crasson, Jean-Louis Lebrave, Jérémy Pedrazzi (ITEM) et Thorsten Ries (Université de Gand) s'engageaient notamment à identifier les problèmes nouveaux soulevés par les traces numériques et à évaluer la façon dont ceux-ci pouvaient à la fois mettre la critique génétique à l'épreuve et l'amener à élaborer de nouveaux concepts et de nouvelles méthodologies.

La figure 1 résume les principales caractéristiques numériques du fonds (hors disquettes).

Les premières investigations

Le 29 novembre 2017, notre équipe (A. Crasson, J.-L. Lebrave, J. Pedrazzi et T. Ries) était reçue à l'Abbaye d'Ardenne par le directeur des collections de l'IMEC qui nous a montré, sans le mettre en marche, un exemplaire du

6. Une empreinte numérique est le résultat d'un calcul sur les données d'un fichier permettant d'obtenir un identifiant de petite taille (quelques octets) qui garantit le caractère unique du contenu.

7. Ou informatique légale : ensemble de procédures d'investigation du contenu de supports numériques préservant entièrement l'intégrité des données originales. Cf. l'article de Jean-Louis Lebrave cité plus haut.

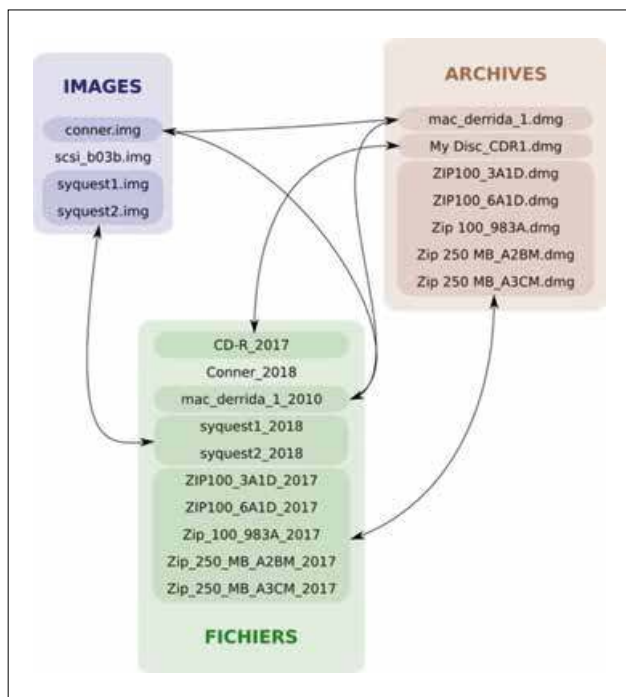


Fig. 1 : Les disques durs et supports externes de mémoire du fonds sous la triple forme d'images forensiques, de disques durs virtuels et de contenus de fichiers. Ce schéma illustre le caractère parcellaire des extractions.

fameux Macintosh Plus 1 MB de la société Apple Computer. Mais ce n'est pas sur la machine elle-même que nous prîmes connaissance de l'ampleur du fonds. Les fichiers extraits du clonage d'une partie des archives numériques avaient été mis à notre disposition – en mode consultation – sur un des moniteurs de la bibliothèque de l'IMEC. Au hasard des circulations dans les arborescences, cette première séance de travail nous a permis de repérer certaines des idiosyncrasies archivistiques de l'auteur, l'abondance des sauvegardes, la richesse de la production textuelle de Derrida, mais aussi bien son art de donner un nom à un fichier ! L'ambition de l'équipe était double : à court terme, présenter au Congrès du cinquantenaire de l'ITEM (octobre 2018) le fonds nativement numérique d'un des plus grands philosophes du xx^e siècle ; à plus long terme, mettre au point un protocole méthodologique et générique pour l'exploitation des données afin de les donner à voir *via* une plateforme et des outils appropriés.

Après cette première visite, l'IMEC a mis à notre disposition en mars 2018 une copie du corpus réalisée conformément aux préconisations de la *computer forensics* : deux disques durs, deux Syquest, cinq Zip. Grâce à *LibreOffice*, suite bureautique libre et gratuite, il a été possible, au gré des pérégrinations dans ce vaste corpus, d'ouvrir et de lire des fichiers produits avec *MacWrite*, le logiciel de traitement de texte en mode graphique apparu dans sa version d'origine avec le premier Macintosh en 1984. Les explorations liminaires en vue du Congrès du cinquantenaire ont permis de cerner les principales difficultés inhérentes au corpus :

- l'absence presque totale de fichiers nativement numériques écrits avec le logiciel Word, alors que les premières expériences d'application de la *computer forensics* à des avant-textes littéraires avaient porté sur ce type de fichiers ;
- la *terra incognita* que représentait l'exploitation des fichiers écrits par Derrida avec le logiciel MacWrite Pro ;
- le foisonnement des fichiers et la dissémination de leur redondance sur tous les supports présents dans le fonds.

En octobre 2018, un premier bilan est présenté au Congrès du Cinquantenaire. Après un panorama des recherches sur l'exploitation génétique de fonds « nativement numériques » par Thorsten Ries⁸, l'équipe a donné les résultats d'une observation plus attentive sur la masse de fichiers du corpus. Le séminaire « Peine de mort » avait été choisi parce qu'il est exhaustif et qu'il est attesté en de nombreux exemplaires sur les différents supports du fonds. Une première analyse forensique (comparaison des empreintes numériques des fichiers, notamment homonymes) nous a permis d'identifier une macro-chronologie d'écriture au sein de l'ensemble des fichiers. En parallèle, l'exploration des systèmes de fichiers a souligné la singularité du système de classement créé par Derrida, elle a mis en valeur son jeu linguistique sur le nommage des fichiers et éclairé la relation d'ordre qu'ils entretiennent entre eux. De manière exploratoire, deux états différents d'un bref fichier de quelques lignes datant de la dernière année de la vie de Derrida ont permis d'identifier des opérations génétiques.

8. http://item-50ans.org/programme/#abst_c2_4

Les développements depuis octobre 2018

La recherche s'est développée sur trois fronts :

Un front codicologique et technique

LES DISQUES DURS

Pour travailler « directement » sur les disques, on peut « monter » une image de ces disques dans les systèmes d'exploitation disponibles (MacOS, Windows, Linux) en spécifiant que leur contenu est en lecture seule, afin de ne pas altérer celui-ci. Une fois montée, l'image est traitée comme n'importe quel autre support de stockage externe. D'une part, on peut lire les fichiers que celui-ci contient si on possède les logiciels avec lesquels ils ont été créés ou si on dispose d'un outil informatique se substituant à ceux-ci. D'autre part et surtout, on peut accéder aux données numériques brutes (sous forme hexadécimale) à l'aide de logiciels spécifiques. Nous avons commencé à exploiter les données propres à chaque fichier (voir *infra*, « Micro-génétique »). En revanche, nous n'avons pas encore tenté de récupérer dans ces images d'éventuels fichiers supprimés par Derrida.

LES DISQUETTES

Le descriptif des disquettes contenues dans le fonds Derrida de l'IMEC reste à faire, et leur exploitation n'a pas encore commencé. Au stade actuel, on sait qu'on pourra lire les formats les plus récents si les disquettes n'ont pas été endommagées, et on espère pouvoir lire les formats plus anciens, mais les phases de test restent à faire. À elles seules, les disquettes constituent en tout cas un formidable matériau génétique par la multiplication des états textuels qu'elles semblent promettre.

À l'époque de Derrida, on sait que les disques durs servaient plutôt d'espaces de travail. Avant la généralisation du courrier électronique, les disquettes constituaient d'une part un moyen commode de transmettre des fichiers à un tiers et de faire voyager les données, et d'autre part un support d'archivage et de stockage à long terme. Tel n'est peut-être pas leur rôle le plus intéressant chez Derrida. Liée semble-t-il à la crainte de perdre des contenus et à un travail nomade, l'utilisation des disquettes par Derrida ne visait probablement pas au premier chef une sauvegarde pérenne

mais la sauvegarde locale, provisoire, d'un travail en cours, ce que paraît confirmer la citation suivante : « On fétichisera, écrivait Derrida, telle disquette archivant une étape d'un *work in progress*⁹ ». Il y a donc tout lieu d'espérer qu'on pourra grâce à elles « dater » des documents en fonction des fichiers qui y coexistent, voire trouver des documents dans des états qui auront été écrasés sur les disques durs.

Un front macro-génétique

Aurèle Crasson s'est penchée sur la notion d'archive chez Derrida et sur son évolution à la suite de l'usage de l'ordinateur. Obsédé par ses archives, Derrida était capable de conserver le moindre bout de papier en se donnant la possibilité d'imaginer qui, après sa mort, le découvrirait en venant consulter l'archive, dans quel dessein, pour quel propos. Il avait consciencieusement choisi ses lieux d'autorité et construit non pas une archive mais un ensemble qui va au-delà d'une mémoire matérielle et qui désignerait plutôt ce qu'il pensait être une « structure survivante des traces ». J. Derrida appartient à une génération confrontée à une transition technologique majeure et la question qui le hante est de savoir comment, à la fin du xx^e siècle, repenser l'archive au regard de l'historiographie, de la mémoire, de la psychanalyse, de la communication et plus spécifiquement des outils informatiques qui la véhiculent. L'archive, expliquait Derrida au cours d'un entretien¹⁰ avec Daniel Ferrer, n'est pas une ; il y a un « processus d'archivage avec des états différents ».

L'usage de l'ordinateur relance chez Derrida la question, récurrente chez lui, de savoir s'il faut élaborer un concept d'archive, technique, politique, éthique et juridique. S'il n'avait pas conscience qu'une archive des réseaux sous la surface aurait une chance de voir le jour, il disait que « le désir de garder, était aussi inséparable du désir de détruire ». Autant détruire (mettre à la corbeille de l'ordinateur) est un geste assumé, autant écraser un fichier en l'enregistrant sous un même nom est une posture tout aussi destructrice mais d'une autre nature ; elle laisse la machine déposséder le scripteur. Lorsque, la veille d'un séminaire,

9. Jacques Derrida, « La machine à traitement de texte », *Papier machine*, Paris, Galilée, 2001, p. 157.

10. Jacques Derrida, « Entre le corps écrivant et l'écriture », Entretien avec Daniel Ferrer, *Genesis*, n° 17, Paris, 2001.

Disque mac_derrida_1_2010. Dossier « animalversaire ». Lecture du fichier « A Mal de Musique » en hexadécimal par le logiciel <i>Hex Fiend</i> .	Les mêmes données lues avec <i>LibreOffice</i>
<pre> 00C00 5175D565 73742D63 65207175 65206C61 00C10 20607573 69717565 3F200046 69637469 00C20 6F6E3A20 20717565 6C7175D5 756E2070 00C30 6F736520 63657474 65207175 65737469 00C40 6F6E20D2 205175D5 6573742D 63652071 00C50 7565206C 61206075 73697175 653FD320 00C60 5175656C 7175D575 6E202061 64726573 00C70 73652063 65747465 20717565 7374696F 00C80 6E2E4461 6E73206C 61206E75 69742ECA 00C90 4C612064 6F6E6E65 20882065 6E74656E 00CA0 64726520 64616E73 206C6120 6E756974 00CB0 2E205361 6E732063 6F6E7465 78746520 00CC0 76697369 626C652E 20C8A5375 7220756E 00CD0 20746F6E 20646F6E 74206F6E 206E6520 00CE0 706F7572 72616974 20706173 20648E63 00CF0 63636363 63636363 63636363 63636363 </pre>	<p>Qu'est-ce que la musique ?</p> <p>Fiction : quelqu'un pose cette question : "Qu'est-ce que la musique ?" Quelqu'un adresse cette question. Dans la nuit. La donne à entendre dans la nuit. Sans contexte visible. Sur un ton dont on ne pourrait pas décider s'il est innocent ou insolent, bien qu'on puisse percevoir</p>
<pre> 00D00 00D10 00D20 00D30 00D40 00D50 00D60 00D70 00D80 00D90 00DA0 00DB0 00DC0 00DD0 00DE0 00DF0 </pre>	<p>240^{ème} caractère du bloc</p> <p>partie du bloc non lue par LibreOffice</p> <p>suite du texte</p>
<pre> 00E00 69646572 2073D569 6C206573 7420696E 00E10 6E6F6365 6E74206F 7520696E 736F6C65 00E20 6E742C20 6269656E 207175D5 6F6E2070 00E30 75697373 65207065 72636576 6F697220 </pre>	<p>...ider s'il est innocent ou insolent, bien qu'on puisse percevoir</p>

Fig. 2 : Jacques Derrida. À gauche, « siliscrit »* de *Qu'est-ce que la musique ?* (les traits violets marquent les délimitations entre blocs**); à droite, résultat de la lecture des mêmes données par *LibreOffice*.

(*) Pour « écrit dans le silicium ». C'est le néologisme que propose Jean-Louis Lebrave en se calquant sur le paradigme existant « manuscrit-tapuscrit ».
(**) Suite élémentaire de 256 – 2 puissance 8 – octets (ou caractères) constituant l'unité de base retenue par les concepteurs de *MacWrite*.

Derrida décide de transformer quelque peu son cours, il ne prend pas la peine d'enregistrer cette transformation sous un autre nom de fichier ou même de la copier sur ses disques de sauvegarde. L'archive est devenue « une et définitivement originale », comme un *Urtext*. Derrida se débarrasse de la masse, du « Mal d'archive » pour n'en garder finalement qu'une colonne vertébrale¹¹.

UNE CARTOGRAPHIE DES SYSTÈMES DE FICHIERS

En utilisant les dates de création et de modification des fichiers ainsi que leurs empreintes numériques, Maëlle Dagot, jeune stagiaire recrutée pour le projet en 2019 grâce à Jean-Gabriel Ganascia (équipe ACASA, LIP6), a créé un graphe des répertoires et fichiers rendant compte de la hiérarchie des arborescences, de l'antériorité des fichiers et de la similarité entre les supports. Une batterie de scripts permet de construire les listes de nœuds et les relations qui les lient entre eux ; l'ensemble est ensuite chargé et exploité dans une base de données.

Un front micro-génétique

Les explications de Laurent Alonso, chercheur à l'INRIA Lorraine et l'un des principaux développeurs d'une brique

logicielle de *LibreOffice*, nous ont permis d'installer la suite d'outils intitulée « *Libmwaw* », pierre angulaire du logiciel *LibreOffice* pour l'importation de fichiers anciens. *Libmwaw* rend lisibles, à partir des données brutes, des fichiers sauvegardés dans un format obsolète, et notamment les fichiers provenant du logiciel *MacWrite*, auquel Jacques Derrida est resté fidèle tout au long de son écriture numérique. Comme nous avons pu le constater immédiatement en comparant la lecture brute d'un fichier sur le disque dur avec celle du même fichier par *LibreOffice*, le logiciel effectue une véritable lecture intelligente des données brutes en effectuant un parcours non linéaire dans celles-ci et en éliminant des incohérences apparentes (fig. 2).

Les clés de cette lecture sont données par les fichiers de débogage utilisés dans les phases de développement

11. Pour Derrida, reprenant le concept de Freud de « compulsion de répétition » et de pulsion de mort, toute archive est entraînée dans ce mouvement de pulsion irréductible de destruction. Dès lors que l'archive est le résultat d'une extériorisation de la mémoire, l'accumulation de documents aurait pour conséquence une totale amnésie. C'est le « Mal d'archive ». L'interprétation est peut-être hasardeuse, mais on peut néanmoins s'interroger sur cet usage qu'il fait de constituer une archive et de détruire ses traces de conservation quasiment dans le même élan en écrasant les versions qui l'ont précédée.

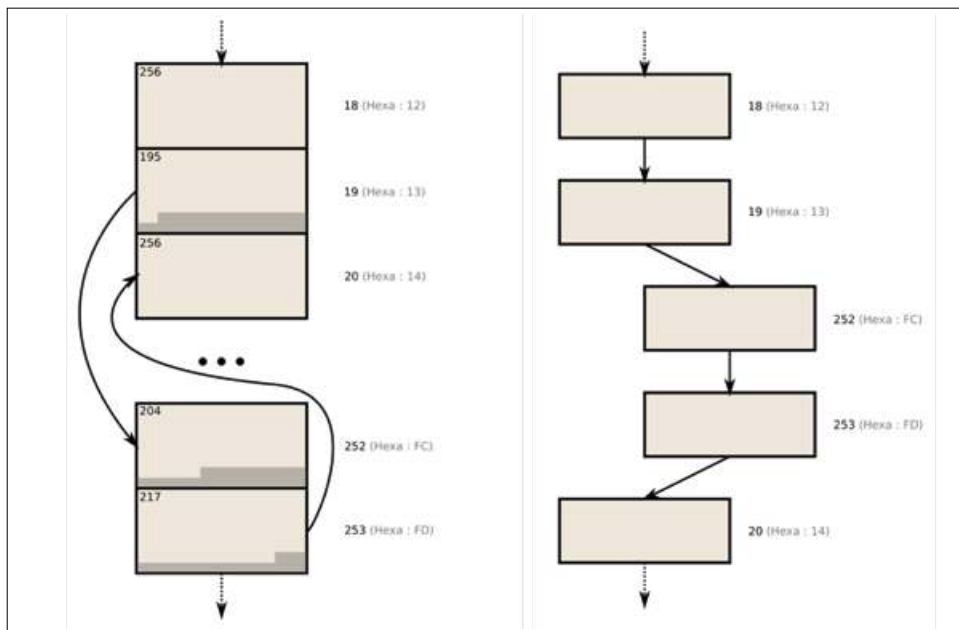


Fig. 3 : Portion du parcours de lecture d'un fichier *MacWrite Pro*.
À gauche, circulation en zigzag dans la succession matérielle des blocs. À droite, le texte lu qui résulte de cette circulation.

du logiciel, fichiers que Laurent Alonso nous a appris à obtenir et à exploiter en recompilant l'outil *Libmwaw* avec une option *debug*.

Libmwaw permet donc de déployer une véritable rétro-ingénierie génétique. Sans entrer dans le détail technique des opérations, le principe de cette ingénierie n'est pas sans rappeler, *mutatis mutandis*, celui que Jean-Louis Lebrave avait identifié au début des années 1980 dans les brouillons manuscrits en distinguant dans ceux-ci des données lisibles et des données d'une autre nature, traces d'opérations d'écriture et de réécriture. L'analogue de cette distinction est ici la présence, au sein même du fichier *MacWrite*, de données textuelles et d'instructions sur la manière de le lire.

Nous n'en sommes qu'aux premiers balbutiements, mais les résultats obtenus sont très prometteurs. En effet, *Libmwaw* appliqué aux fichiers *MacWrite Pro 1.5* (les plus nombreux) fournit les informations suivantes : 1) l'unité informatique¹² d'écriture dans *MacWrite* est le bloc de 256 caractères ; 2) l'ordre de lecture des blocs n'est pas nécessairement séquentiel ; *Libmwaw* permet de restituer cet ordre ; 3) un bloc peut ne pas être lu dans son intégralité physique de 256 caractères ; les fichiers générés par *Libmwaw* donnent la longueur utile du bloc. Longueur *utile*, car celle-ci n'est pas toujours de 256 caractères. Lorsque c'est le cas, cette donnée paraît être corrélée avec une instruction de lecture sautant un ou plusieurs blocs pour revenir ensuite en arrière. Par exemple, dans le même fichier « A MaL de Musique », le bloc 18, de longueur utile 256, est suivi du bloc 19, de longueur

utile 195, donc inférieure à 256. La lecture s'interrompt au caractère 195 du bloc 19, saute 232 blocs en allant au bloc 252 suivi du bloc 253, pour revenir ensuite au bloc 20. L'interprétation génétique de ces données reste à affiner, mais on peut d'ores et déjà faire l'hypothèse qu'elles correspondent à une addition à l'intérieur du bloc 19 (fig. 3).

Il apparaît qu'on peut raisonnablement faire l'hypothèse que ce bouleversement de la suite linéaire des blocs résulte d'une opération de réécriture. En effet, le stockage des données sur un support numérique ne connaît pas l'équivalent d'un becquet ou d'une correction interlinéaire : toute intervention sur du déjà écrit sera nécessairement stockée dans une partie du disque encore non écrite, située au-delà des emplacements déjà utilisés, ou libérée par effacement de tous les caractères d'un bloc. Si on admet que le bloc de 256 caractères est métaphoriquement l'équivalent d'une page manuscrite¹³, la page de silicium ne comporte ni marges ni espaces interlinéaires susceptibles d'accueillir des réécritures. C'est un bloc compact à l'intérieur duquel seules des suppressions peuvent apparaître sous la forme de l'instruction « ne pas lire au-delà du caractère 195 ». L'indication de la longueur du bloc rapportée à 256 paraît

12. Informatique : en effet, l'utilisateur Derrida n'avait aucun moyen de savoir qu'il franchissait une frontière d'unité tous les 256 (28) caractères.

13. Avec la réserve, capitale, évoquée dans la note précédente : il s'agit d'une unité de traitement par l'ordinateur, et non d'une unité perceptible par le scripteur.

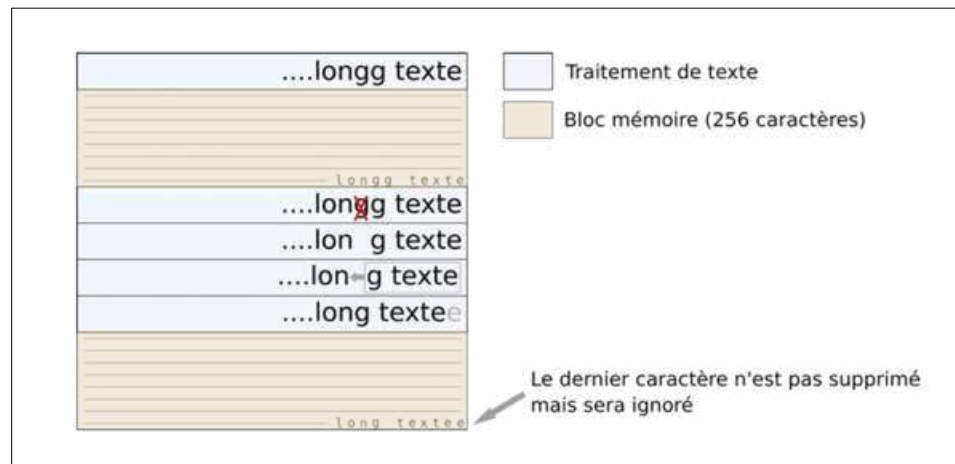


Fig. 4 : Incidence d'une suppression dans le traitement de texte sur un bloc mémoire.

donc être l'équivalent d'une marque de réécriture dans un manuscrit (fig. 4).

Nous sommes loin d'avoir élucidé le détail des instructions de « non-lecture » et de celles de lecture en zigzags à l'intérieur des blocs, et ce sera le prochain objectif de nos recherches. Mais nous pouvons déjà identifier des suppressions, comme c'est le cas dans l'exemple précédent.

Perspectives

Septembre 2019 marque une nouvelle étape dans le travail d'élaboration d'une codicologie informatique à partir de l'observation des traces de réécriture conservées dans la structure informatique interne des ordinateurs de J. Derrida. Le projet dorénavant intitulé « Derrida Hexadécimal »¹⁴ se déploie autour de deux axes :

– Le premier est génétique ; il s'agit tout d'abord de constituer une archive inédite et d'élaborer un dossier génétique à partir des traces numériques. Il conviendra ensuite d'analyser l'ensemble des fichiers, y compris ceux exhumés en les comparant aux archives papier du philosophe et d'explorer la façon dont le traitement de texte a interagi avec la créativité verbale de Derrida (néologismes, concepts, lapsus, lexicque d'enregistrement des fichiers, arborescence informatique, etc.). Au détour de cette exploration, on cherchera à comprendre comment s'est forgée et transformée la notion d'archive chez Jacques Derrida au cours du passage de l'écriture manuscrite à l'écriture sur ordinateur.

– Le second axe vise le développement de systèmes de circulation au sein d'archives numériques hétérogènes et non indexées (visualisation de données) : il s'agit d'une part d'adapter les dispositifs d'édition génétique élaborés

à partir des dossiers papier au foisonnement des données numériques, et d'autre part d'exploiter, et si nécessaire, d'améliorer les outils d'analyse et de comparaison de textes, comme *Médite*, pour analyser les fichiers.

Dans le détail, nos objectifs sont les suivants :

Lecture des supports

LES DISQUES DURS RÉTIFS

Sur trois disques durs provenant des trois ordinateurs déposés à l'IMEC, deux n'ont pas pu être lus pour l'instant : le premier pour un problème d'alimentation, et le second en raison d'un problème sur les têtes de lecture. Le contrat EUR nous permettra de financer leur traitement par une société spécialisée afin qu'elle prenne en charge la réalisation d'une copie forensique. À défaut de pouvoir exploiter les deux ordinateurs disparus dont nous connaissons l'existence grâce au témoignage de Derrida, nous devrions donc disposer de deux images-disques supplémentaires qui élargiront notre champ d'investigation et rendront les données génétiques nativement numériques moins lacunaires.

LES DISQUETTES

On réalisera une image numérique de chaque disquette, en associant à celle-ci une photo des deux faces de la disquette afin de conserver la trace des procédures d'archivage et de nommage du philosophe. Pour ce faire, une machine dédiée a été préparée. Le transfert est prévu à partir de

14. Il bénéficie d'un financement de l'EUR (École universitaire de recherche), PSL « Recherche-formation *Translitterae* ».

l'automne 2019, mais reste tributaire des opérations de description. Les données seront ensuite intégrées au corpus.

Exploitation macro-génétique

Grâce à l'exploitation du travail réalisé par Maëlle Dagot, on peut envisager une exploration systématique de l'arborescence des fichiers, et par là même la constitution, à partir du foisonnement anarchique des fichiers dans les supports, d'équivalents numériques des *dossiers génétiques*. Une des questions soulevées sera de soupeser l'incidence de cette équivalence sur la définition même du dossier génétique.

Exploitation micro-génétique

Comme on l'a rappelé plus haut, les outils de la *Libmaw* permettent, *via* un mode « débogage », d'exploiter les fichiers temporaires générés lors de la transformation des fichiers en documents *LibreOffice*. Ces fichiers contiennent une description parfois très lisible des informations encodées dans le fichier source. Les formats anciens de fichiers texte n'écrasant pas systématiquement les espaces mémoire libérés, nous pensons pouvoir identifier des opérations génétiques grâce aux structures mises à nu de ces fichiers.

Nous commencerons cette « rétro-ingénierie génétique » avec l'identification d'opérations d'écriture dans les textes écrits avec *MacWrite Pro* et *MacWrite II*. À partir des fichiers de débogage, nous pourrions proposer un affichage des fichiers mettant en évidence les parties « à ignorer » et rechercher à quelle opération de réécriture (addition, suppression ou remplacement) elles correspondent. Le jeu apparemment très complexe entre la longueur utile des blocs et leur place dans la séquence de lecture devrait aussi nous fournir des informations sur le second composant minimal des remplacements, c'est-à-dire les additions.

À titre indicatif, signalons une donnée intéressante que nous commençons seulement à exploiter. Les formats *MacWrite* sont de très loin les plus représentés, ils constituent

par exemple 98 % des fichiers dans les images forensiques. Malgré la position dominante prise rapidement par Microsoft Word sur le marché des logiciels de traitement de texte, Derrida, comme on l'a signalé plus haut, est resté très fidèle aux logiciels Apple intégrés aux versions des premiers Macintosh. Il convient toutefois d'être prudent : on ne connaît pas pour l'instant le contenu des deux disques durs « rétifs », ni celui des quelque 500 disquettes que contient le fonds.

Conclusion

Avec l'approche forensique, c'est une nouvelle dimension qui s'offre à la critique génétique. Elle interroge à nouveau la théorie génétique en confrontant ses lois fondamentales et sa méthodologie à un support jusque-là inexploré.

En outre, elle permet d'envisager dans un avenir proche plusieurs types de productions :

– une étude sur les caractéristiques de l'invention scripturale propre à Derrida – notamment l'invention linguistique (jeux de signifiant), et l'élaboration des concepts, recherche qui demeure l'un des principaux enjeux de ce projet ;

– plus généralement, un manuel de « génétique numérique/forensique » destiné à tous ceux qui souhaitent exploiter des archives numériques ; on y présentera les avancées théoriques liées au projet, on les illustrera d'exemples rencontrés dans l'exploration des données du corpus Derrida, et on décrira les outils et comment les utiliser ;

– enfin, dans le prolongement du travail de Maëlle Dagot, un dispositif en ligne d'exploration enrichie des arborescences de fichiers : circulation à l'intérieur des différents supports ; visualisation des relations typées – inclusion, parenté, identité, antériorité, postériorité –, en exploitant des graphes génétiques ; ce type d'outils est indispensable pour les fonds numériques, si l'on veut à la fois optimiser la conservation archivistique des données et faciliter leur exploitation pour une recherche génétique.