Comment sortir les textes du musée computationnel ?

Pour un autre scénario de l'édition numérique (la table de montage)

Christine Noille

Professeur en Rhétorique et théorie littéraire
Littérature du XVIIe siècle, Sorbonne Université (UMR CELLF)

Résumé : Numériser des données et documenter les bases participent d'un processus global de patrimonialisation des textes sous forme de données : mais la patrimonialisation passe ici par l'éviction de la textualité. La *distant reading* n'est que métaphoriquement une lecture car les machines ne lisent pas, au sens littéral. D'où l'urgence, pour les communautés littéraires, fédérées autour des activités de *close reading*, d'évaluer les limites du paradigme éditorial numérique actuel et d'envisager une conversion des humanités numériques à des problématiques de *data-visualization* susceptibles de répondre à leur attente en matière de lecture augmentée.

Mots-clés : numérisation – édition – *distant reading* – *close reading* – *data-visualization*

En guise d'introduction, nous commencerons par un aveu et un mode d'emploi.

L'aveu, c'est que la tâche que nous nous sommes fixée – réfléchir aux liens entre patrimoine et humanités numériques à partir du cas emblématique de l'édition de corpus textuels massifs – a été compliquée par la diversité de ses destinataires et la pluralité de ses buts.

Diversité des destinataires tout d'abord. Pour une part en effet, nous parlons entre littéraires de l'édition de corpus textuels, essayant d'en théoriser les procédures : mais notre description a vocation à être compréhensible par des non-spécialistes en ce qu'elle s'adresse également aux informaticiens et aux linguistes. Et pour une part, nous parlons des principes et des enjeux de la numérisation tels qu'ils sont débattus

par les linguistes et les informaticiens : mais nous tenterons là aussi de les redécrire pour qu'ils soient accessibles aux littéraires.

Pluralité des buts ensuite. Notre propos vise en effet à interagir sur ces deux communautés relativement étrangères l'une à l'autre que sont donc les *philologues* d'une part, en les mettant en garde sur ce que veut dire éditer, numériser, patrimonialiser un texte ; et les *informaticiens* et *linguistes* d'autre part, en leur soumettant un certain nombre de critiques et de requêtes concernant les protocoles de numérisation actuels et les besoins en termes d'évolution des systèmes.

Où l'on voit que le propos sera truffé de récriminations et d'objurgations. Le mode d'emploi en sera par conséquent très simple : que chacun lise ce qu'il pourra, en faisant comme s'il n'était pas concerné par les critiques... Car « le meilleur parti que le lecteur ait à prendre est de se mettre d'abord dans l'esprit qu'il n'y a aucune de ces [réflexions] qui le regarde en particulier, et qu'il en est seul excepté, bien qu'elles paraissent générales ; après cela, je lui réponds qu'il sera le premier à y souscrire. » (La Rochefoucauld, 1664)

1 Patrimonialiser

L'idée de patrimoine est liée à celle de perte avérée ou de disparition potentielle. Elle légitime par là même des opérations de conservation et d'exposition systématiques (Babelon et Chastel, 1980). Mais s'il importe de muséaliser l'objet patrimonial, c'est qu'il est crucial pour une société donnée de le sauver, d'en garder la mémoire. En effet, l'idée de patrimoine est tout autant liée à celle d'identité culturelle (Arpin, 2000). En amont se sont donc développées des stratégies dites de patrimonialisation, entendues comme processus de valorisation culturelle et identitaire d'un objet donné (ou d'une pratique donnée : objet immatériel). En aval, la muséalisation entraîne un dispositif très particulier de conservation : le musée ne conserve pas en l'état l'objet de culture, mais à part (Malraux, 1951, Déotte, 1986). Parce qu'ils ont été valorisés comme éléments identitaires en danger, ils sont séparés de leur contexte pour être conservés, exposés, étudiés comme documents représentatifs (Desvallées et Mairesse, 2010).

Importer la notion de patrimoine pour penser notre rapport à des textes a alors un certain nombre de conséquences, idéologiques et méthodologiques, car la patrimonialisation est loin d'être une opération neutre.

- Elle postule tout d'abord qu'au niveau des institutions culturelles il existe quelque chose comme une valorisation des corpus textuels, qui en fait des objets suffisamment identitaires pour qu'ils participent d'un patrimoine immatériel donné, que ce soit à titre d'échantillon (de témoignage d'une culture) ou à titre de monument (de référence pour une culture), que ce soit par exemplification ou par exemplarité.

- Elle postule également que la pensée de la perte soit présente dans notre rapport au texte, qu'il s'agisse d'une perte matérielle (fonds d'archive en péril) ou d'une perte symbolique (textes significatifs d'un *moment* culturel et tombant ou tombés dans l'oubli).

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

- Elle implique enfin une opération d'extraction et de réification de la textualité, en l'engageant en dehors de son territoire d'usages et de pratiques, dans des protocoles de pérennisation et d'accessibilité.

D'où une interrogation tout à fait décisive, avant même de poser les questions de support et de méthode : avons-nous envie ou besoin de patrimonialiser un texte ? Pouvons-nous dans un même geste à la fois promouvoir sa valeur et stigmatiser son inactualité ? Voulons-nous œuvrer à sa mémoire (en le conservant à part) plus qu'à son actualisation (en en réinventant un usage possible) ? Ce sont là des questionnements absolument décisifs qui, par-delà les positions individuelles contingentes que les uns ou les autres pouvons prendre sur le sujet, engagent de fait le rapport général qu'une culture, que notre culture, entretient avec ses textes.

2 Lire

L'analyse a souvent été faite (Genette 1966, Charles 1984, Noille 2015) que l'on peut distinguer dans l'histoire de la culture occidentale entre deux grands styles de rapport aux textes fondateurs (quelles que soient les dénominations qui les désignent : modèles, autorités, littérature…), qui ont alterné et parfois se sont superposés. Il y a en effet une différence – épistémologique, méthodologique, idéologique – entre *interpréter* un texte (l'insérer dans un réseau de notes ou de commentaires) et *analyser* un texte (ou comme on dit par facilité, « décrire » quelque chose d'un texte).

Différence épistémologique tout d'abord, entre deux finalités de savoir étrangères l'une à l'autre, une investigation historique ou idéologique sur le sens littéral et sur sa tension vers d'autres niveaux de sens et vers le monde de référence d'une part ; un savoir sur l'architexte d'autre part, autrement dit sur les procédures, les formes et les modes de fonctionnement des œuvres de langage.

Différence et même opposition de méthode, ensuite, entre deux discours critiques, avec d'un côté un discours d'éclaircissement et de reprise herméneutique du sens, en symbiose commentatrice (citationnelle) avec le fil du texte, et de l'autre côté un discours d'analyse et de mise à distance (Genette, 1966). La séparation se fait ici sur les manières d'étudier les textes, entre deux positionnements de l'auctorialité critique, le pathos de l'herméneute d'une part, son engagement (furieux ou enthousiaste, selon les époques) dans la poursuite du sens ; l'*ithos* du théoricien d'autre part, entre désengagement et domination.

Différence idéologique enfin et surtout, quant aux présupposés sur le statut et la définition du texte qui légitiment et motivent ces deux gestes critiques (Charles, 1984). Car de fait, la différence dans les méthodologies de relecture des textes engage et implique deux conceptions pour le coup contradictoires concernant le rapport au texte : d'un côté, l'idée de l'autorité et de la monumentalité du texte détermine une enquête qui en déploie toujours plus les richesses et la singularité ; de l'autre, l'idée que le texte est le composé, la combinatoire – plus ou moins arbitraire, plus ou moins contingente –, de structures et de formes reproductibles et modifiables s'accorde avec une pratique manipulatrice et interventionniste, qui opère une sérialisation toujours possible du texte sur d'autres corpus exemplifiant des formes comparables, sur d'autres versions

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

exemplifiant des formes modifiées. Ou, si l'on veut, une culture du texte – que Michel Charles nomme « commentatrice » – contre une culture qu'il nomme « rhétorique » et dont la fonction est de « casser » l'autorité du texte, de le considérer comme un assemblage de possibilités, « un instrument à fabriquer d'autres discours » (Charles, 1984).

Bilan de cet état de l'art, nous avons donc deux rapports aux textes fondateurs et deux conceptions de la textualité hétérogènes qui ont parcouru notre histoire : d'un côté il y a l'annotation philologique et l'interprétation (Moyen-Âge, époque contemporaine) et de l'autre l'analyse des formes (antiquité, cultures d'Ancien Régime), comme il y a d'un côté un engagement dans le sens et de l'autre une technique de description, d'un côté l'annotation commentatrice et de l'autre l'analyse, d'un côté une investigation en profondeur dans le texte et de l'autre un positionnement en surplomb, d'un côté une sanctification et une sanctuarisation de la textualité et de l'autre un démontage jubilatoire (Noille, 2015).

Arrêtons-nous un instant sur les conséquences épistémologiques fondamentales qui en découlent. Quelle que soit en effet la relation que la culture noue avec des textes de référence, qu'elle les considère comme fondateurs ou comme modèles, il s'avère que c'est précisément sur cette relation que portent de façon essentielle nos recherches en littérature.

Qu'est-ce à dire ? Au moins deux choses : non seulement nos disciplines ont un rapport fort avec le paradigme du *texte* (du latin *textus*, tissu, qu'il soit discursif ou narratif, factuel ou fictionnel) – c'est-à-dire avec tout ensemble d'énoncés dont la composition relève d'un tissage, d'une *écriture*, et dont la réception relève d'une appropriation compréhensive qui soit à la fois mise en ordre et mise en sens, préhension et compréhension. Mais par là même, nos travaux et nos enseignements mettent l'accent sur la fécondité de la *lecture*, de toute lecture en tant que s'y engagent à la fois et inextricablement un travail cognitif, une expérience affective et une *praxis*. Et nous œuvrons en particulier à la légitimité de la *relecture*, en tant que s'y construit un *savoir-lire* (Noille, 2014) qui est absolument identitaire dans nos disciplines – dans le sens où elles ont partie liée avec l'idée d'une lecture productive, d'une activité poïétique d'actualisation et de re-création des textes (Schuerewegen, 2012) ; avec l'idée d'un travail sur la textualité (ou sur la discursivité, sur la narrativité, sur la fictionnalité…) qui s'apprenne et se développe, qui soit un champ de recherches tout autant qu'un secteur de l'enseignement.

Il apparaît alors clairement qu'entre le texte et la lecture s'intercale un geste absolument séminal, celui de l'édition, dont la tâche première est de délimiter son corpus. Or, la notion de *corpus textuels* recouvre de fait au moins trois choses :

- des écrits qui n'ont pas été et ne sont toujours pas considérés comme des textes de référence : en termes de volumétrie, la masse de loin la plus importante ; il s'agit d'énoncés conservés sur divers supports (manuscrits ou imprimés) et participant d'une relation dédramatisée à la textualité ; ils sont les échantillons de pratiques culturelles passées ou présentes ; leur édition participe d'une démarche épistémologique d'historien (documenter un fonds) ;

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

- des textes modélisés (exemplaires) ouvrant sur des pratiques multiples de (re)production ; leur édition participe d'une démarche pédagogique d'éducation aux formes par les exemples, d'initiation à l'écriture par la lecture guidée et les morceaux choisis.

- des textes cardinaux dédiés à la pratique de la lecture littéraire ; leur édition participe de l'entretien de leur monumentalisation.

Quel que soit le rapport au texte, l'édition est donc un objectif de choix, que l'on s'inscrive dans un programme documentaliste, dans une perspective de productivité discursive, dans un projet de valorisation ; que l'on cultive le texte comme données documentaires, comme échantillon de formes, comme lieu identitaire. Dans tous les cas, il est pertinent de les transmettre et de les publier, de les introduire dans l'actualité des débats et lectures, bref, de les rendre accessibles. Mais ce faisant, l'édition est loin alors de recouper les postulats et les implications propres à la patrimonialisation.

C'est à ce point qu'il convient de regarder les choses d'un peu près.

3 Éditer

Pourquoi éditer ? Pour mettre à jour : l'étymologie même nous le rappelle, qui fait de l'*editor* celui qui rend public, qui fait de la publicité par voie de publication. Nous dirons qu'est inscrit dans tout programme d'édition un imaginaire fort de l'actualisation, du don fait au texte d'une nouvelle actualité par la grâce du support matériel qui assure dans le *hic et nunc* du présent sa mise au (grand) jour.

L'édition restitue au texte une présence et un présent, là où la patrimonialisation participe d'un archivage (du passé comme passé) et d'une séparation. Faire entrer les textes dans le patrimoine c'est autant les adouber (en reconnaître une efficience culturelle datée) que les aliéner en les objectivant, en les projetant dans la distance de leur inactualité. Éditer un texte au contraire, c'est faire le pari que le présent sera riche des pratiques qu'il saura en faire et c'est aller un peu plus loin que le pari, c'est en proposer des pratiques possibles, par le biais des interventions sur le texte (modernisation de l'orthographe et de la ponctuation, mise en page, index, tables, annotations. . .).

Comme on peut alors s'en douter, l'enjeu est moins vif dans le cas où les corpus textuels ont un intérêt documentaire. Mais à supposer que l'édition participe d'une culture de l'exemple ou d'une culture littéraire, l'on peut concevoir qu'elle devienne littéralement séminale, dans la mesure où les pratiques de formalisation et d'imitation (sous l'Ancien Régime) comme les pratiques de commentaire et d'interprétation (au XX^{e} siècle) passent par l'actualité du texte, par sa présentation dans un format lisible.

Et l'on retrouve ici les intentions épistémologiques qui président dans nos communautés interprétatives au travail d'édition scientifique :

- l'édition vise à produire un savoir du texte et un savoir sur le texte. Deux finalités peuvent être ici distinguées : la finalité d'information (fournir un état du texte praticable par la communauté scientifique, pour sa documentation et son travail : l'intégrer dans le stock commun des références interrogeables et récupérables) ; la finalité ency-

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

clopédique (compléter ou couvrir systématiquement un corpus, une chaîne textuelle, qu'il s'agisse d'un ensemble historique, géographique, auctorial, générique…).

- l'édition vise à produire un discours d'accompagnement (commentaire), à introduire une focale idéologique à propos du texte et autour du texte. Là encore, on peut reconnaître à l'œuvre deux finalités : une finalité de valorisation (autoriser un texte, le légitimer comme modèle dans une culture rhétorique ou comme œuvre de référence dans une culture de la littérature) ; une finalité identitaire (souder une « communauté interprétative » autour du travail d'édition et de valorisation du texte).

- enfin, l'édition vise donner une actualité à un texte, autrement dit à susciter sa lecture. Permettre à un texte de trouver, d'élargir, de renouveler son lectorat, ouvrir la lettre du texte au plaisir de la lecture en continu, cursive et intégrale, générer un désir du texte, qui soit autant désir de relecture que désir d'appropriation, désir d'immersion que désir de re-création, est à l'origine même de l'activité d'édition. Ils légitiment à ce titre des opérations complémentaires d'actualisation du texte (modernisation de l'orthographe, de la syntaxe, de la ponctuation… ; travail typographique au service du sens : intervention sur les caractères, les paragraphes, les pages… ; travail sur l'intelligibilité du dispositif global : titulature, tables, index…).

C'est à ce point (et à ce point seulement, du point de vue qui nous occupe ici, analytique et non pas institutionnel) que se pose avec pertinence la question des supports éditoriaux. En effet, on peut opérer sans trop d'anachronismes la restriction suivante : le geste philologique au XX^{e} siècle a porté sur des corpus ayant à voir avec la littérature (que ce soit de façon linéaire ou plus dialectique, on y revient à l'instant) et pour lesquels la publication avait valeur de consécration, d'introduction dans le champ de pratiques dédiées. Or, l'un des apports décisifs de la critique littéraire au XX^{e} siècle est de n'avoir cessé de retravailler à l'extension (idéologique) du littéraire et à l'ouverture (méthodologique) du concept de texte.

D'une part en effet, la littérature a été pluralisée (les littératures) et dé-normée, intégrant à elle tout candidat pressenti, c'est-à-dire tout texte dont un commentaire récapitule la présence (l'actualité) : en particulier les textes non réédités, parfois en raison de leur rareté, ou de leur échec, parfois aussi en raison de leur volume (romans pastoraux et précieux, mémoires, sommes…). D'autre part, la textualité elle-même a vu son mode d'existence devenir éminemment problématique, dans la mesure où le texte n'est plus cet espace clos enserré entre la première et la dernière ligne, mais est désormais ouvert sur des co-textuels multiples, réseau intertextuel, variantes, versions concurrentes, manuscrits et traces de sa genèse, généalogie de ses reprises (Charles, 1984).

Où l'on commence à voir ce qui fait problème : la difficulté qu'il y a à manipuler cette masse inflationniste de textes n'est pas épistémologique – car prendre en compte les différents états de la textualité ou des textes atypiques oblige (simplement) à compliquer le geste critique (Charles, 1984). La difficulté est phénoménologique et accessoirement commerciale : comment enserrer une telle masse textuelle dans les limites d'un support imprimé ? L'éditeur se heurte à un certain nombre d'insuffisances liées à la matérialité du livre :

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

- insuffisance en termes de longueur, les conditions du marché excluant la multiplication des volumes, de même que des ouvrages trop volumineux ;

- insuffisance en termes de manipulation, la succession des pages rendant malaisée la collation des variantes et plus encore des versions concurrentes ;

- insuffisance en termes de mise en page, le cadre focal ne permettant pas, dans le cas de l'édition d'un manuscrit par exemple, la confrontation sans hiérarchie des transcriptions multiples (diplomatique, linéarisée, chronologisée...).

Et l'on retrouve alors les communautés qui sont aujourd'hui les plus actives dans le domaine de l'édition numérique : médiévistes (œuvres à branches multiples), éditeurs des grands romans et des collections de chroniques, spécialistes de génétique textuelle. Ainsi mise en perspective, l'édition numérique n'est par conséquent ni un effet de mode ni un moyen d'intégrer des circuits d'institutionnalisation : elle rémunère les défauts du support papier ; elle propose une actualisation et une présence possibles pour des textualités que le livre imprimé trahirait parce que sous-dimensionné. Ou, ce qui revient au même, elle est un pas vers le rêve et le besoin d'un *livre augmenté*.

4 Numériser

L'ennui, c'est qu'une édition numérique n'est pas du tout un livre augmenté.

La numérisation d'un texte peut être d'abord en mode image, comme on le sait, et restituer par composition pdf (Portable Document Format) le feuilleté d'un *opus* imprimé. L'usage est tenu pour équivalent à celui de l'objet imprimé, soit qu'on utilise l'écran pour faire défiler les pages, soit qu'on les imprime. La limite est bien sûr dans le traitement informatique de la donnée : le texte est converti sur la toile mondiale en un stock d'images (au même titre qu'un carnet de voyage photographique ou un catalogue en ligne de produits, par exemple), insensible à toute reconnaissance optique de caractères (OCR, Optical Character Recognition).

Mais l'édition numérique va bien au-delà de cette numérisation, qu'elle l'intègre ou qu'elle l'évite. Elle consiste en effet à ce que ce ne soient pas (que) les pages matérielles numérisées qui constituent les données informatiques, mais d'abord les caractères et les mots du texte. Opérer la saisie « au kilomètre » de la lettre du texte dans un logiciel de traitement de texte *ad hoc* est alors une étape intermédiaire, désormais purement technique car prise en charge par un outillage informatique (OCR, logiciel de correction...). Le travail éditorial proprement dit consistera à enrichir la saisie « au kilomètre » de la lettre du texte par trois opérations :

a) l'« annotation sémantique » ou insertion automatique de balises descriptives autour de mots ou de séquences de mots (avec, en amont, une réflexion sur le choix des balises pertinentes) ;

b) la répartition en unités numériques : espaces ou « pages » html dont l'identité numérique (par exemple l'URL : localisateur uniforme de ressources) est stable et dont la description est opérée (par les « métadonnées » ou indications sur la ressource encapsulée) ;

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

c) l'insertion de liens hypertextuels permettant de mettre en relation dynamique un espace numérique avec un autre, une base de données avec une autre.

Si l'on reprend plus en détail ces opérations :

a) La balise fonctionne comme une description de la forme et du fond des énoncés (Lebarbé et Meynard, 2011), comme une étiquette pouvant indiquer la catégorie graphique (rature, suscription...), grammaticale (morphologie, syntaxe), fonctionnelle (titre, paragraphe, chapitre ...) voire stylistique (nom propre, citation...) permettant d'identifier un mot ou une séquence de mots par les attributs qu'il exemplifie. La règle est ici l'objectivité (la positivité) de l'attribut en question – un terme *est* ou *n'est pas* un dérivé de *util**, un substantif *est* ou n'*est pas* un nom propre, un groupe nominal *est* ou *n'est pas* un titre, une séquence *est* ou *n'est pas* une citation – ce qui revient à éliminer tout attribut dont la reconnaissance n'est pas automatique mais dépend de l'analyse et de la relation critique, autrement dit toutes les qualités *plus ou moins* exemplifiées (Leriche, 2006). Et il est clair par exemple qu'avec l'attribut de la citation (Compagnon, 1979 ; Genette, 1982), on atteint rapidement les limites de l'annotation automatique.

b) La « page » numérique fonctionne en amont comme un espace sémantique informé, doté lui aussi d'une étiquette décrivant la forme et le fond de son contenu informationnel ; et en aval comme un espace visuel structuré. En amont tout d'abord, la description de l'ensemble des informations contenu dans la page passe donc par la définition de « métadonnées » pertinentes, c'est-à-dire par le remplissage d'une étiquette donnant des indications sur la nature du document, lesquelles, pour le domaine éditorial, sont souvent adaptées à partir des entrées utilisées dans le référencement bibliographique standard (items des date, lieu, auteur, etc.). Le but est de permettre le catalogage de la page (et partant son « moissonnage »). En aval, la construction visuelle relève d'une interface graphique qui obéisse aux principes de communication issus du Web design. Son référent est l'écran (de l'ordinateur) et en ce sens, le terme d'« image » ou de « plan » (cinématographique) lui conviendrait assurément plus que celui de « page ». Les composantes y sont organisées selon un scénario codifié (onglets, cadres formatant les ressources focales, etc.).

c) Si le plan de la « page » est à la fois stabilisé et autonome, il est également ouvert sur un processus de consultation dynamique grâce auquel un plan numérique (ou plus précisément un « nœud », une unité d'information contenu dans la page) est en relation (sémantiquement pertinente) avec un certain nombre d'autres nœuds contenus dans d'autres plans – offrant alors une efflorescence de parcours d'investigation que les spécialistes du Web nomment hypertextualité. Dans les faits, l'ouverture hypertextuelle dans le domaine de l'édition numérique est étroitement modérée ; soit qu'elle s'effectue sans basculement sur une page (informations complémentaires ouvertes dans des fenêtres temporaires ou « pop-up ») ; soit qu'elle donne accès à un base documentaire d'appui (élucidation des références) ; soit qu'elle opère un basculement repéré sur un lieu précisé du « rail » textuel (« page » suivante, précédente ou localisée dans une table des matières) ; soit enfin qu'elle autorise une collation semi-automatisée des différents modes de numérisation d'une séquence donnée (mode image, variantes en mode texte...).

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

La réalisation de ces opérations passe donc par l'encodage (par exemple en XML-TEI) et par l'usage des « ontologies » informatiques. L'édition devient ainsi une archive et une *knowledge base*, dont les données formalisées sont interrogeables et manipulables par la machine. Une telle entreprise aide à alors comprendre pourquoi les humanités numériques sont porteuses d'une transformation des pratiques savantes (Berra, 2012) : car si le domaine du numérique abonde en métaphores issues de la pratique du livre et de la lecture (page, parcours, hypertexte…), on aura sans doute commencé à comprendre qu'il ne s'agit là que de métaphores.

Nous ferons ici un pas de plus et nous dirons alors : l'édition numérique n'est pas même un livre. Ce sera là du moins notre hypothèse heuristique (et politique : une prise de conscience et un appel, si l'on veut).

5 Ne plus lire / *Distant reading*

Quels sont tout d'abord les usages que visent les opérations décrites ci-dessus, de balisage, de catalogage, d'insertion d'hyperliens ?

Les métadonnées visent à permettre un catalogage automatique de la page par des moteurs de recherche ; le balisage vise à permettre une requête automatique sur l'ensemble des données ou ressources (l'ensemble des mots et séquences encapsulés dans les balises) ; l'hyperlien vise à nourrir un scénario de visite. Autrement dit, tandis que l'hyperlien, discret dans l'édition numérique, nourrit tout un fantasme du parcours libre et de la sérendipité lectoriale, le travail lourd de marquage systématique (des données et des métadonnées) est là pour répondre à un traitement automatique du corpus (extraction et indexation des données). Historiquement, son élaboration informatique a correspondu aux demandes de la linguistique de corpus, laquelle établit un lien épistémologique fort entre l'interrogeabilité automatique et la constitution de corpus de travail.

Ce qui veut dire que l'édition numérique est conçue pour être « traitée » par des machines (et interprétée par des linguistes), et non pas pour être lue par des humains (et relue par des critiques). Ou, pour reprendre le titre d'un article déjà ancien (Clement, Steger, Unsworth, Uszkalo, 2008), « How Not to Read a Million Books », comment faire avec un million ou plutôt désormais avec trente millions de livres numérisés, qui ne sont pas à lire parce qu'ils ne sont pas faits pour la lecture, autrement dit, parce qu'ils ne sont pas lisibles ? Pour nous littéraires qui nous approprions les outils et les terminologies d'opérations éditoriales numériques co-élaborées par les linguistes et les informaticiens, nous nous sommes laissé prendre au triple piège de la *systématicité*, de l'*accessibilité* et de l'*interrogeabilité* :

- Un corpus numérisé systématiquement n'est pas une édition intégrale, il est un stock organisé d'unités d'information.

- De même, un texte accessible n'est pas un texte lisible, il est un texte dont les protocoles de numérisation permettent l'interopérabilité entre les systèmes informatiques et l'indexation automatique.

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

- De même encore, un texte interrogeable n'est pas un texte offert au feuilletage et à la confrontation des passages ; il est un réseau de données extractibles sur lequel peut opérer un moteur de recherche.

On pourrait cependant remettre en cause l'usage qui est fait ici de la notion de lisibilité et se demander si, en surface, le texte qui s'affiche à l'écran n'est pas tout simplement aussi lisible que n'importe quel texte imprimé, pendant que dans l'arrière-boutique, le texte numérisé aurait également été pensé pour être manipulable par des opérateurs artificiels. Deux arguments viennent détruire cette belle illusion.

Un argument de fait tout d'abord : chacun a pu faire l'expérience qu'une lecture suivie d'un énoncé long mis en ligne en mode texte n'est pas praticable, d'une part pour des raisons de confort et de fatigue visuelle, d'autre part pour des raisons d'intelligibilité des « masses » textuelles, de repérage dans la progression – et ce quels que soient les scénarios mis en place pour lever l'obstacle (voir par exemple l'usage des résumés dans l'édition numérique d'*Artamène* : Bourqui et Gefen, 2002-2009). De fait, le lecteur ne peut pas maîtriser et structurer le cursus de sa lecture (à la différence des ouvrages numérisés en mode image ou pdf ainsi que des livres numériques sur tablette), ce qui revient à dire qu'il n'arrive pas à mettre en forme le sémantisme qu'il déchiffre (Noille, 2013). Dans une tentative de lecture en ligne, on est toujours un peu perdu, et rapidement tenté de se replier sur une simple posture de « visiteur » (terme consacré qui dit bien ce dont il est ici question) – à moins d'en revenir à la version pdf ou/et imprimable de la page numérique, ce qui revient à se replier sur une pratique de l'imprimé (ou son équivalent en mode image).

Un autre argument vient, au niveau des principes, corroborer le fait que le basculement d'un texte sur un support numérique a perdu en route quelque chose de sa textualité. La description de la forme et du contenu des séquences textuelles et des « pages » numériques a pour enjeu de transformer l'énoncé en information, ou plus exactement chaque énoncé en une ou plusieurs informations qui lui sont corrélées (l'ensemble des métadonnées qu'il exemplifie). Autrement dit, le texte numérisé n'est plus un *textus*, un *tissu* tramé de mots et de sens, il est devenu *data*, collection de données autonomes. Ce n'est pas un texte qui défile, mais une base de données qui est mise en scène d'un plan séquence à un autre. La numérisation d'un texte est le protocole par lequel la textualité est transformée en cellules d'information.

L'imprimé me donne un texte à lire, le numérique me donne une base de données à consulter. Que vais-je bien pouvoir devenir, moi qui suis critique littéraire, professionnelle de la lecture, maniaque de la relecture ? Coopérative et aventureuse comme je le suis, je saurai m'adapter. Je vais arrêter de lire, et réfléchir aux requêtes intelligentes que je pourrais soumettre aux moteurs de recherche des éditions numériques : de façon à associer l'interrogation quantitative et statistique dont l'outil m'offre la possibilité, à des problématiques qui fassent sens sur des corpus de masse que je n'aurais pas lus, qui en permettent le profilage et la cartographie automatiques (par exemple en les sérialisant par thème, forme, genre, période ou continent…). Les corpus constitués par le résultat des requêtes automatiques seraient ensuite interprétés à la lumière des connaissances que les recensions manuelles avaient jusqu'ici effectuées. Et l'on arriverait enfin à la constatation que le traitement à distance de corpus massifs re-

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

nouvellerait l'histoire littéraire (puisque c'est bien d'elle qu'il s'agit) et ferait émerger un autre *story-telling* de nos genres, formes, thèmes et périodes. Tout cela non seulement sans avoir lu, mais précisément parce qu'on n'aurait pas lu, parce qu'on aurait su faire autre chose.

Trêve de plaisanterie : comme on le sait, la chose existe déjà, brillamment théorisée par Franco Moretti (2008 et 2013). Elle a pour nom « *distant reading* » (opération métaphoriquement dite de « lecture » opérant « à distance » du texte) et offre à la recherche contemporaine sur les corpus de textes une des perspectives les plus stimulantes et les plus imaginatives. Mais il convient peut-être de resituer sa portée : en valorisant les données organisées et archivées par l'édition numérique, elle continue par d'autres moyens à muséaliser un corpus, en le séparant de ses pratiques actuelles (pratiques de lecture cursive et intégrale) et en le prenant comme objet culturel global à documenter et à cataloguer. La numérisation enferme le texte dans le musée computationnel où sont stockées les *big data* : la *distant reading* en opère un décryptage de masse.

Dès lors, est-il besoin d'ajouter que l'historiographie assistée par ordinateur n'est évidemment pas une *distant reading*, ni même une *machine reading*, une « lecture » algorithmique (Hayles, 2012) : elle est d'abord une *not reading*, une non-lecture (Mueller, 2007), en ce qu'elle exclut tous les usages de la lecture, amatrice ou professionnelle, empirique ou méthodique ; ou encore, en ce qu'elle démode toute pratique du texte comme survivance corporatiste (et perte de temps).

Bref, lire moins et quantifier plus (Moretti, 2000), voilà le nouveau dogme des historiens de la culture.

6 Relire, malgré tout. Pour un autre scénario de l'édition numérique (la table de montage)

Nous pouvons à présent reprendre les fils de notre investigation.

Il apparaît tout d'abord que les travaux d'histoire culturelle répondent à une pragmatique muséale et à des problématiques qui lui sont subordonnées, de conservation et de documentation d'un fonds ; ce qui revient à dire que l'interprétation historiographique du traitement informatique des corpus est dans la stricte continuité d'un archivage des textes par leur conversion numérique. Numériser des données et documenter (interroger, interpréter) la base participent d'un processus global de patrimonialisation des textes sous forme de données : mais la patrimonialisation passe ici par l'éviction de la textualité. Car les machines ne lisent pas, au sens littéral, pas plus que ne sont destinées à être lues, interprétées ou actualisées les données informées mises en scène dans les plans séquences de nos écrans. Nous sommes donc ici aux antipodes du « livre augmenté », en train de prendre la mesure de nouveaux territoires, ceux des subjectivités computationnelles et de l'accès algorithmique aux données de l'information, à l'intérieur d'un horizon culturel également bouleversé, où les textes littéraires ont cessé d'être actifs pour relever de stratégies de sauvegarde et d'exposition, de catalogage et de muséalisation.

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

Et pourquoi pas en effet, s'il est avéré que la lecture n'a plus d'actualité dans le champ de nos pratiques ? si notre époque et nos disciplines de savoir ont mieux à faire que de lire ? si l'invention du futur passe par la renégociation de notre rapport aux discours ?... Sauf que rien n'est clair au pays des signes et que décider du sens de notre histoire relève au mieux d'un pari idéologique, au pire d'un *story-telling* à visée politique. Peut-être sommes-nous en train de congédier la littérature et le rapport particulier qui a été entretenu sous son nom avec les textes. Peut-être nous tournons-nous vers une culture qui est en passe de faire prévaloir la production des discours sur l'identification de ses références. Peut-être même entrons-nous dans un moment où développer l'intelligence des données est plus nécessaire que lire du texte ou produire du discours pour rendre notre monde habitable. Mais plutôt que d'envisager les évolutions sous l'angle du changement de paradigme, il est sans doute possible et assurément plus efficace de plaider pour des styles divers de pensée du texte qui coexistent et parfois se recoupent, de façon à ne pas entrer dans le piège interprétatif et d'en revenir aux méthodologies de l'édition et/ou du numérique.

En effet, nous pouvons tout à fait acter l'existant et nous réjouir que des protocoles de numérisation textuelle enrichie permettent de convertir des corpus textuels massifs en données interopérables et interrogeables – et que des méthodologies d'interprétation culturelle permettent aux requêtes computationnelles de faire sens. Mais pourquoi abandonner aux nouveaux venus le terrain de la légitimité et de l'importance qu'ont les travaux que nous menons ? Ne sont-ils pas partie prenante dans la façon dont notre culture redéfinit constamment son rapport à la langue, au discours, à l'écrit et à la créativité verbale (Schuerewegen, 2012) ? Et en quoi un moissonnage statistique serait-il plus objectif et partant plus décisif qu'une analyse de texte, alors que tous deux sont également adossés à des protocoles herméneutiques qui leur donnent sens, à des discours d'interprétation (ici culturelle et là littéraire), à des manipulations discursives aussi complexes qu'humaines ? Faudrait-il même les opposer, les hiérarchiser, quand l'un et l'autre mettent le prix de leur démarche dans l'inventivité et la complexité des hypothèses, dans l'ingéniosité et la cohérence des systèmes échafaudés pour rendre compte de l'investigation, qu'elle soit requête ou analyse ?

Quoi qu'il en soit, les investigations opérées sur les corpus massifs numérisés ne sont pas sans intérêt, loin de là, et documentent notre connaissance de l'histoire culturelle. Mais l'intérêt qu'il y a à savoir faire sens avec les mots des autres et à savoir mettre en formes nos propres mots est tout aussi urgent, tant au point de vue épistémologique que politique. Nous redirons ici, avec humilité mais aussi avec fierté, ce que nous rappelions plus haut, à savoir que notre savoir est d'abord un savoir-lire ; que nos travaux et nos cours servent à construire des lectures sur les textes, à en transmettre le savoir-faire, à en donner le plaisir ; et que notre discipline trouve son identité et sa valeur dans un mythe, au sens le plus noble, à savoir que mieux on lit et plus on lit (plus on écrit), qu'il y a un itinéraire et une ascèse dans la compétence lectoriale, et que la relecture, tout comme la littérature, est elle aussi – au bout du chemin – une « aventure d'être » (Richard, 1954).

Littéraires, nous sommes des lecteurs et nous construisons un savoir sur et de nos lectures : est-ce une raison suffisante pour nous cantonner à l'arrière-garde des révo-

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

lutions technologiques ? En tant qu'éditeurs de textes, nous sommes les premiers (les seuls, peut-être) à avoir éprouvé les insuffisances de l'imprimé, à les avoir recensées, analysées, théorisées : et nous avons méthodiquement rêvé d'un livre augmenté.

Alors nous y voilà. Posons pour une fois les bonnes questions, nous, les communautés philologues, aux informaticiens et linguistes qui règlent notre accès à la numérisation.

Pourquoi ne serait-il pas possible d'ajouter d'autres protocoles et d'inventer d'autres interfaces qui augmentent la potentialité de *lecture* et permettent une pratique lectoriale enrichie de ces mêmes corpus ? Pourquoi nos disciplines fondées sur la relecture professionnelle ne seraient-elles pas à leur tour prescriptives à l'égard des communautés informaticiennes afin qu'ensemble elles élaborent un modèle pour penser une donnée informatique complexe, le film du texte ?

Car soit que notre rapport à la textualité participe d'une monumentalisation ou qu'il se tourne vers l'inventivité discursive, il n'en reste pas moins tributaire de ce que les archéologues du champ nomment la « *close reading* » et qu'il est plus juste de nommer tout simplement « *reading* », arpentage du territoire textuel en vue d'en approfondir le paysage ou d'en appréhender la cartographie. Dans les deux cas en effet, la lecture « de près » n'est en rien myope ni focalisée sur le détail (encore moins étroite ou fermée…), elle est simplement tissage des énoncés dans une dynamique de compréhension opérée par un cerveau humain et non par une machine computationnelle. Et s'il est intéressant de voir émerger de nouvelles recherches dédiées aux systèmes numériques de données textuelles, il n'en reste pas moins que le projet d'*édition* (au sens philologique de mise à jour) reste d'actualité, qui vise à décliner au présent un texte dans la pratique d'une lecture humaine (faut-il dire : *human reading* ?) ; tout comme reste aussi sensible et urgent le besoin d'un « texte augmenté » qui autorise une édition des textualités envahissantes (par leur longueur, par leur multiplication, par leur complexité…).

Une première réponse possible est de distinguer d'une part le travail de numérisation enrichie en mode texte, soumis aux protocoles de balisage et d'indexation mis en place par et pour le traitement automatique des corpus ; et d'autre part le travail de mise en scène des données disponibles dans un ou plusieurs scénarios de visite et de consultation, grâce à l'élaboration d'interfaces graphiques innovantes ainsi que des outils de conversion en mode image (générateurs de pdf ou de documents à imprimer). Le texte serait toujours converti en *data*, mais les *data* elles-mêmes pourraient être configurées en unités lisibles de texte, en quelque sorte ; et le cursus lectorial serait mimé dans la navigation hypertextuelle. C'est à peu près l'état où nous en sommes aujourd'hui dans l'édition numérique, avec d'un côté de la barrière, des travaux d'équipe dédiés à la mise au point de projets de documentation culturelle rentabilisant la consultation automatisée des corpus « dormants » de *data* textuelles (voir *The Stanford Literary Lab*, 2010 – en cours [1]) ; et de l'autre côté des programmes destinés à modéliser les scé-

1. https://litlab.stanford.edu/ [consulté le 17 juin 2020]. Page d'accueil : « The Stanford Literary Lab is a research collective that applies computational criticism, in all its forms, to the study of literature. » (« Le laboratoire littéraire de Stanford est un collectif de recherche qui applique la recherche informatique, sous toutes ses formes, à l'étude de la littérature. »

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

narios graphiques pour l'exploration humaine des corpus numérisés (voir la structure fédérative interne MÉTILDE, 2010-2014 [2]).

Une autre réponse, plus imaginative, serait possible. Il convient de la formuler en guise de conclusion ou plutôt de sollicitation. En superposant les outils de consultation dynamique et les outils de programmation de montage filmique [3], l'édition numérique gagnerait à garantir une lisibilité augmentée des textes en proposant à la consultation la possibilité de manipuler la textualité selon six modes opératoires fondamentaux dans les opérations de relecture critique, dont les deux premiers sont bien évidemment consubstantiels au dispositif de l'imprimé, mais dont les suivants en énumèrent précisément les faiblesses :

1. un mode de structuration séquentielle de la donnée textuelle : où l'élément minimal ne serait pas le mot, l'expression ou le format étiquetés selon des catégories grammaticales ou stylistiques (titre, paragraphe, chapitre, citation...), mais une extension syntagmatique prédéfinie, la séquence, dont l'étiquette sémantique serait d'être une unité de lecture à combiner (l'équivalent numérique de la page : comprise non pas comme contenant visuel, mais comme contenu intelligible, tissage sémantique local et transitoire d'énoncés) ;

2. un mode de défilement ordonné des séquences, avec possibilité de retours, de reprises, ou de sauts (voir la visualisation d'un film...) ;

3. un mode de consultation en parallèle de différentes séquences textuelles avec convocation possibles de leurs co-textes amont et aval : autrement dit, une consultation en parallèle de différents films textuels (voir les murs d'écrans des salles de montage télévisuel) ;

4. un mode de convocation duel des séquences parallélisables : soit selon un principe de choix (séquences en autonomie) ; soit selon une interrelation renseignée (collation intelligente avec séquences hyperliées : projet Schola-rhetorica.fr, 2012-en cours) ;

5. un mode de sélection des séquences et de mise en réseau ; avec possibilité de variation sur le nombre des séquences et sur l'ordre du rail (voir les outils de montage des séquences vidéo) ;

6. un mode de réécriture des séquences et d'écriture de séquences supplémentaires susceptibles d'être interpolées n'importe où dans le rail du film textuel (insertion de commentaires : voir les outils d'annotation et de surlignage sur tablettes).

Le livre augmenté libérerait la table de travail en la désencombrant des multiples ouvrages et papiers qui composent l'environnement réel du critique. Il proposerait

2. La structure fédérative interne MÉTILDE (Manuscrits, Éditions, Technologies de l'Information, Linguistique et Document Électronique), projet mené entre 2010 et 2014 par Cécile Meynard, Thomas Lebarbé et Viviane Clavier au sein des équipes Traverses, LIDILEM et GRESEC à l'université Stendhal, Grenoble pour chercher de nouvelles façons de présenter, de consulter, d'analyser et d'interpréter l'objet manuscrit. Voir quelques-unes de ces expérimentations sur le « Stend'lab », http://stendhal.demarre-shs.fr/stendhal-wordpress/?page_id=161 [consulté le 17 juin 2020]

3. Voir les recherches en modélisation géométrique, analyse d'image, informatique graphique et vision par ordinateur du Département « Géométrie et image » du Laboratoire Jean Kuntzmann (UMR 5224, UGA/CNRS), où un programme de modélisation intuitive pour les environnements narratifs a été initié par l'équipe IMAGINE.

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

sur le bureau virtuel de l'ordinateur des outils pour mener une lecture augmentée, fortement interventionniste. Il ferait du lecteur le monteur réalisateur de son parcours d'appropriation et de la relecture-écriture le modèle actif de l'exploration numérique.

C'est ainsi que les données numériques fondamentales dans l'édition de corpus seraient à tisser et non pas uniquement à interroger ; à composer et non pas uniquement à empiler ; à actualiser et non pas uniquement à archiver ; à investir et non pas uniquement à encapsuler ; à lire et non pas uniquement à compter.

En attendant, avec les outils qui sont les nôtres d'annotation et d'indexation propres au traitement automatique des corpus, quand nous faisons des éditions numériques, ne nous trompons sur ce que nous faisons : un travail de conservateur, destiné à nourrir les enquêtes statistiques des documentalistes. Prenons soin malgré tout de les enrichir par des scénarios de consultation intuitifs et productifs. Et réservons, faute de livre augmenté, le travail d'édition scientifique à l'imprimé, ce mode définitivement restreint et contraint d'actualisation de la textualité.

Bibliographie

Arpin, Roland *et al.* (2000) *Notre Patrimoine, un présent du passé*, Québec. http://bel.uqtr.ca/1239/1/6-19-681-20060117-1.pdf [consulté le 10 juin 2020].

Babelon, Jean-Pierre ; Chastel, André (1980) « La notion de Patrimoine », *La Revue de l'Art.*

Berra, Aurélien (2012) « Faire des humanités numériques », dans *Read/Write Book 2. Une introduction aux humanités numériques*, P. (dir.), OpenEdition. http://books.openedition.org/oep/238?lang=fr [consulté le 10 juin 2020].

Charles, Michel (1984) *L'Arbre et la source*, Paris, Seuil.

Clement, Tanya ; Steger, Sara ; Unsworth, John ; Uszkalo, Kirsten (2008) « How Not to Read a Million Books », URL : http://people.brandeis.edu/ unsworth/hownot2read.html [consulté le 10 mars 2015].

Compagnon, Antoine (1979) *La seconde main ou le travail de la citation*, Paris, Seuil.

Desvallées, André ; Mairesse, François (dir.) (2010) *Concepts clés de muséologie*, Paris, Armand Colin.

Deotte, Jean-Louis (1986) « Suspendre – Oublier », revue *50, Rue de Varenne* No. 2, p. 29-36.

Genette Gérard (1976) [1re éd. 1966] « Structuralisme et critique littéraire », *Figures I*, Paris, Seuil, coll. « Points », p. 159-161.

Genette Gérard (1982) *Palimpsestes. La littérature au second degré*, Paris, Seuil.

Hayles, Katherine (2012) *How We Think : Digital Media and Contemporary Technogenesis,* Chicago, University of Chicago Press.

La Rochefoucauld François de (1664) « Avis au Lecteur » dans *Réflexions ou Sentences et Maximes morales*, Paris, Claude Barbin.

Leriche, Françoise (2006) « Quel balisage pour les corpus épistolaires numériques ? De l'annotation traditionnelle du « document » à une analyse générique et pragmaticienne » dans *Corpus en lettres et sciences sociales*, Rastier, F. et Ballabriga M. (dir.,), Paris, Texto. http://www.revue-texto.net/Parutions/Livres-E/Albi-2006/Sommaire.html [consulté le 19 juin 2020].

Malraux, André (1951) *Les voix du silence. Le musée imaginaire*, Paris, NRF.

Meynard Cécile ; Lebarbé, Thomas (2011) « Au croisement des lettres, de la linguistique et de l'informatique : Les Manuscrits de Stendhal en ligne », *Fabula-LhT*, No. 8, « Le partage des disciplines ». http://www.fabula.org/lht/8/lebarbe.html [consulté le 19 juin 2020].

Moretti, Franco (2000) « Conjectures on World Literature », *New Left Review*, 1.

Moretti, Franco (2008) *Graphes, cartes et arbres : Modèles abstraits pour une autre histoire de la littérature*, Paris, Les Prairies ordinaires.

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843

Moretti, Franco (2013) *Distant Reading*, New York, Verso.

Mueller, Martin (2007) « Notes towards a user manual of MONK ». https://apps.lis.uiuc.edu/wiki/display/MONK/Notes+towards+a+user+manual+of+Monk [consulté le 19 juin 2020].

Noille, Christine (2013) « Le discours d'un Prince (*Énéide*, I, 597-610). Archéologie de la disposition », *Exercices de rhétorique* [En ligne], 2 - 2013, mis en ligne le 17 octobre 2013. http ://rhetorique.revues.org/176 [consulté le 19 juin 2020].

Noille, Christine (2014) « Le plaisir des textes (et nous) » dans *Un autre dix-septième siècle. Mélanges en l'honneur de Jean Serroy*, Noille, Ch. et Roukhomovsky, B. (dir.), Paris, Champion.

Noille, Christine (2015) « La forme du texte : rhétorique et/ou interprétation », *Fabula-LhT*, No. 14, « Pourquoi l'interprétation ? ». http://www.fabula.org/lht/14/noille.html [consulté le 20 juin 2020]

Programme : Projet « Artamene.org », dir. scientifique Claude Bourqui, conception technique Alexandre Gefen. http://www.artamene.org/projet.php (2002-2009) [consulté le 18 mai 2019].

Programme : Structure fédérative interne « MÉTILDE » : « Manuscrits, éditions, traitement de l'information, linguistique et document électronique », université Grenoble Alpes, E.A Litt&Arts (anciennement Travers 19-21), E.A. Lidilem, E.A. Gresec (2010-2014).

Programme : Projet « Schola-Rhetorica.fr – Portail de rhétorique », dir. Christine Noille, Sorbonne-Université – RARE Rhétorique de l'Antiquité à la Révolution, E.A. Litt&Arts (2012-en cours).

Programme : « The Stanford Literary Lab », dir. Franco Moretti. http://litlab.stanford.edu (2010-en cours) [consulté le 20 juin 2020].

Richard, Jean-Pierre (1981) [1re éd. 1954] « Avant-propos » dans *Littérature et sensation*, Seuil, réédition dans *Stendhal Flaubert*, Seuil, « Points ».

Schuerewegen, Franc (2012) « Brèves remarques sur Franco Moretti », *Poétique* 169, p. 85-96

Published in : Meynard C. ; Lebarbé T. ; Costa S. (dir.) (2020), *Patrimoine et Humanités numériques*, Editions des archives contemporaines, ISBN : 9782813003843, doi : https ://doi.org/10.17184/eac.9782813003843