



HAL
open science

Numérique

Jean-Baptiste Camps

► **To cite this version:**

Jean-Baptiste Camps. Numérique. En quête de sources: Dictionnaire critique, 2021, pp.282-287.
halshs-03048437

HAL Id: halshs-03048437

<https://shs.hal.science/halshs-03048437>

Submitted on 14 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Numérique

VOIR AUSSI: BASE DE DONNÉES, COMPUTATION, DONNÉE, MÉTADONNÉE, TEXTE NUMÉRIQUE.

Sans revenir en détail ici sur l'histoire longue de l'intérêt des historiens pour les sources numériques et les données chiffrées, on s'attachera surtout au sens plus restreint de *numérique*, qui renvoie à la fois au code binaire de l'informatique et aux réalisations qui s'appuient sur celle-ci, qu'il s'agisse du traitement, codage et diffusion des sources ou de leur analyse.

Dans l'histoire, désormais longue, des relations entre l'historien et l'ordinateur, il importe de distinguer l'utilisation de ce dernier pour des tâches lourdes et répétitives telles que l'indexation des sources, s'accompagnant, plus récemment, de leur diffusion plus aisée (*TEXTE NUMÉRIQUE), d'un usage plus analytique ou herméneutique, allouant au calcul un rôle central dans la démarche d'analyse (*COMPUTATION), en cherchant notamment à identifier des régularités difficilement perceptibles par l'œil humain. Ce second aspect, à savoir l'exploitation des sources par le calcul, a été pendant longtemps et est peut-être encore le parent pauvre des « humanités numériques », comme cela a de nombreuses fois été déploré, depuis la période d'essor initial (Milic, 1966; Froger, 1968) jusqu'à encore une date assez récente (Busa, 2004).

Bien sûr, l'utilisation du calcul comme instrument pour la recherche historique et philologique ne démarre pas avec l'apparition des ordinateurs à la fin des années 1940. La constitution des disciplines philologiques et historiques en sciences au cours du XIXe siècle donne lieu, dans la seconde partie de ce siècle, à l'apparition de la stylométrie, à l'élaboration d'une première 'machine à compter' les occurrences textuelles, de même qu'au développement de l'histoire économique et sociale, autant d'approches qui connaissent des prolongements nouveaux au cours de la première moitié du XXe siècle (*COMPUTATION). Toutefois, l'apparition des premiers ordinateurs a été un phénomène décisif, de par sa recontre précoce tant avec la linguistique qu'avec les tendances historiographiques d'alors. Les applications au langage et aux textes ont très tôt engendré un fort intérêt. La traduction automatique, après la Deuxième Guerre mondiale, est peut-être le premier problème d'importance que l'on imaginait pouvoir rapidement résoudre, entraînant engouement et attentes très hautes, notamment pour ses applications stratégiques dans un contexte de Guerre froide, avant le coup d'arrêt brutal du rapport ALPAC en 1966 (Léon, 2001; Hutchins, 2001). Concomitamment, le premier projet impliquant des sources anciennes est vraisemblablement l'*Index Thomisticus* de Roberto Busa (*TEXTE NUMÉRIQUE). Cette «automatisation» des recherches philologiques et littéraires que Dom Froger juge «courante» en 1968, tient alors surtout à l'indexation, mais pas seulement: de premiers logiciels de collation, de stemmatologie ou d'aide à la critique conjecturale voient le jour (Froger, 1968).

Chez les historiens français, les premiers travaux ayant recours à l'ordinateur apparaissent

à la toute fin des années 1950 et durant les années 1960, dans les *Annales*, qu'il s'agisse des travaux de Michel Mollat, Jean Delumeau ou Marcel Couturier (Genet, 1986), ou bien encore d'Emmanuel Le Roy Ladurie et François Furet (Chaunu, 1972). Sous l'influence d'Ernest Labrousse et de Fernand Braudel, les historiens quantitativistes des décennies suivant la Deuxième Guerre mondiale trouvent dans l'ordinateur (ou la « machine mécanographique ») un outil pour assembler des données, trier et quantifier de l'information. C'est donc naturellement qu'ils réfléchissent, tant d'un point de vue théorique que technique, à la manière d'extraire des sources et de structurer l'information utile: nature des documents, méthodes de dépouillement, structure des fiches et saisie ou codage (Daumard et Furet, 1959; Couturier, 1966). Chez Delumeau, par exemple, il s'agit à partir de sources portant sur les escales des terre-neuviens à Saint-Malo de 1681 à 1690, de saisir (ou plutôt de faire saisir par les opératrices) 3000 fiches descriptives, d'où sont ensuite extraits des éléments tels que le nombre et tonnage annuel des navires, le tonnage moyen, les ports d'attache, la proportion de retours directs ou indirects, etc. (Delumeau, 1961). Cela n'exclut pas toutefois de premières tentatives plus analytiques, à l'instar des travaux de Gardin & Garelli (1961), s'appuyant sur la théorie des graphes pour analyser des données extraites de tablettes cunéiformes (*COMPUTATION). Aux États-Unis, la même période voit aussi des travaux pionniers tels que ceux des Bailyn en 1959 sur les armateurs du Massachussetts en 1697-1714 ou dans la décennie suivante de Rabb sur l'expansion commerciale anglaise au XVIIe siècle, ou bien encore de Herlihy sur le *catasto* florentin (Shorter, 1971).

La décennie voit aussi l'émergence de revues dédiées aux applications de l'ordinateur dans les sciences humaines, qu'il s'agisse de la *Revue informatique et statistique dans les sciences humaines* (1965), qui fait suite à la création du Laboratoire d'Analyse Statistique des Langues Anciennes (LASLA) à l'Université de Liège en 1961 ou bien encore de *Computers and the Humanities* à New York en 1967. À Strasbourg puis Nancy, l'action de Paul Imbs mène à la naissance du *Trésor de la langue française* en 1960, parmi les tous premiers laboratoires de sciences humaines doté d'un ordinateur (un Bull-Gamma 60), dont proviendront à la fois un dictionnaire et une base textuelle, *Frantext* (Robert, 2010). Toujours à Strasbourg, ce même Paul Imbs convainc Charles Muller, un des fondateurs en France de la statistique lexicale, d'entreprendre une thèse. D'autres centres voient le jour, et on citera sans exhaustivité aucune, le comité « Calculateurs », dirigé par Marcel Couturier à l'École des hautes études (Lemny, 2017), ou encore le Centre d'analyse documentaire pour l'archéologie au CNRS par Jean-Claude Gardin à la fin des années 1950 (Bourin et Zadora-Rio, 2013), comptant dans ses rangs archéologues, mathématiciens et informaticiens. Le monde anglo-saxon n'est pas en reste avec le *Literary and Linguistic Computing Centre of the University of Cambridge* (1964) ou l'*Historical Data Center* à Princeton vers 1971 (Shorter, 1971). En Allemagne, le centre de calcul de l'Université de Tübingen recrute en 1966 le jeune Wilhelm Ott comme chercheur en application de l'ordinateur aux sciences humaines (EADH, s.d.). C'est dans ce contexte de renouvellement méthodologique et technique qu'Emmanuel Le Roy Ladurie peut déclarer en mai 1968 que l'« **historien de demain sera programmeur ou ne sera pas** » (Le

Roy Ladurie, 1968).

Dès lors, les formations à la programmation pour les sciences humaines font l'objet de débats et sont mises en place dans certaines institutions (Milic, 1966), à l'instar de la formation au « *Nichtnumerische Datenverarbeitung* » donnée par le *Deutsches Rechenzentrum Darmstadt* en 1966 (EADH, s.d.). La période est aussi à la publication de manuels (Dearing, 1967; Heller et Logemann, 1967; Leed, 1966; Müller, 1968; Shorter, 1971). Malgré tout, à l'époque comme aujourd'hui, un des points d'étranglements demeure la formation des historiens. L'enthousiasme des premières générations se heurte au manque de formation et peut-être à la mauvaise volonté, d'une partie de la communauté. Outre la difficulté technique, certains aspects nécessaires à l'analyse computationnelle pouvaient ainsi être en décalage avec la pratique historique plus traditionnelle: l'ordinateur demande d'explicitier les implicites, les présupposés ou les hypothèses (Furet, 1971), a notamment des difficultés avec les notions floues, et nécessite une pensée plus modulaire et des énoncés plus pragmatiques (Milic, 1966). En outre, tant la saisie que l'opération encore complexe de ces machines contraint à accepter la dimension nécessairement collaborative de telles entreprises (Bullough, 1967).

Partie de l'étude des données économiques, l'histoire quantitative s'étend aux données démographiques puis, à partir des années 1970, à l'analyse des « systèmes de civilisation » (Chaunu, 1972), lorgnant alors vers l'analyse de corpus textuels, et la sémantique ou lexicographie quantitatives. Pionnier de ce passage des historiens de l'analyse de séries chiffrées à celle de textes, le travail mené par Gilbert Shapiro sur les cahiers de doléance (Shorter, 1971) ou l'ouvrage de Régine Robin, *Histoire et linguistique* (Robin, 1973; cf. Bourin et Zadora-Rio, 2013). Se plaçant dans le cadre plus général du *linguistic turn*, la lexicométrie se développe en histoire dans les années 1970 et encore 1980 (Bourin et Zadora-Rio, 2013), avec notamment, pour la période médiévale, les travaux de Jean-Philippe Genêt, d'Alain Guerreau et de leurs élèves, ou, pour la période contemporaine, les travaux d'Antoine Prost (par ex. Prost, 1974). Certains historiens rejoignent ainsi les recherches en lexicométrie et en analyse du discours, initiés tout d'abord à Strasbourg par des chercheurs comme Paul Imbs, Charles Muller et ses élèves, ainsi qu'à Nice avec les travaux de Pierre Guiraud ou Étienne Brunet, ou l'ENS de Saint-Cloud et l'Université de Nanterre (Cinquin, 2011), qui donnent lieu à l'émergence d'un courant d'analyse de données textuelles à la française, pas sans liens avec l'école française d'analyse des données et les outils développés notamment par Jean-Paul Benzécri (tels que l'analyse factorielle).

Toutefois, après la floraison initiale de la période 1950-1970, l'informatique en sciences humaines entre progressivement dans une période qu'on a pu qualifier de consolidation (Hockey, 2004), voire, à partir du milieu des années 1970, de recul, d'abord léger mais qui s'accroît nettement au cours des deux décennies suivantes, particulièrement chez les historiens français. Au cours des années 1970, l'informatisation rapide des pratiques historiennes tarde à se produire, de même que la démocratisation de l'accès ou de la disponibilité des ordinateurs. Des verrous technologiques (concernant notamment la

structuration des données), demeurent, tandis que les historiens intéressés à l'usage de l'ordinateur continuent de construire leur communauté. On peut faire l'hypothèse que les historiens se sont retrouvés face à deux chemins, devant lesquels ils hésitent encore aujourd'hui: d'un côté la transformation irrémédiable de leur discipline d'origine (telle qu'elle s'est produite par exemple dans les sciences du vivant: qui songerait à faire de la biologie 'non numérique'?) ou la constitution en discipline autonome de l'histoire numérique ou de la philologie computationnelle, allant de pair avec la poursuite du paradigme d'origine (que l'on songe au statut de la linguistique informatique et du traitement automatique de la langue vis-à-vis de leur discipline mère).

Les évolutions historiographiques participent de ce recul, avec un basculement vers des approches moins sociales et plus individualistes ou micro-historiques, vers des théories de l'histoire comme récit, de même que l'éloignement des cadres que pouvaient fournir tant le marxisme que le structuralisme, le tout allant de pair avec une remise en cause du rôle de la quantification dans l'établissement de la scientificité de la démarche historique, du « vertige de l'exhaustivité » et du « vertige statistique » par des auteurs comme G. Béaur ou B. Lepetit (Bourin et Zadora-Rio, 2013; Zalc et Lemercier, 2008). On pourrait aussi être tenté de voir dans ce relatif hiver du numérique en histoire, la conséquence d'un choc entre la hauteur des attentes engendrées par l'apparition de l'informatique et la déception devant la grandeur des défis posés par l'approche computationnelle, aussi bien d'un point de vue pratique que théorique. On pourrait alors tracer un parallèle avec les cycles vécus par des champs tels que la traduction automatique, ou plus largement encore l'intelligence artificielle, alternant période d'exaltation et hivers causés par déceptions et coupes brutales des financements (Hutchins, 2001; Crevier, 1993).

Dans l'attente d'une synthèse quantifiée, portant sur les pratiques numériques des historiens durant cette période et qui prenne en compte les différences que l'on peut soupçonner par champ historique, aire géographique, il est nécessaire de rester nuancé et prudent. Hockey, qui attribue le label de période de consolidation aux années 1970-1980 (Hockey, 2004), la décrit comme marquée par les progrès de la diffusion des méthodes élaborées à la période précédente, la fondation de *l'Association for Literary and Linguistic Computing* lors d'une réunion au King's College London en 1973, la mise en place de conférences régulières au Royaume-Uni puis également aux USA, qui a bientôt sa propre *Association for Computers and the Humanities* (1978). Dans le monde anglo-saxon, l'heure est ainsi au développement des centres de *Humanities Computing* (Hockey, 2004), tandis que Belgique et Pays-Bas notamment ont eux-aussi moins vécu cette période de glaciation, avec du côté des sciences du texte, les grandes entreprises de Dees et de ses élèves à Amsterdam, tant en dialectométrie qu'en stemmatologie (*COMPUTATION), ou bien l'existence d'une chaire de philologie computationnelle et d'un Centre Informatique de Philosophie et de Lettres à Liège (Delcourt, 1984). En France, et du côté des revues, le *Médiéviste et l'ordinateur* est fondé en 1979 à l'IRHT, et *Histoire et Mesure* en 1985. Nuancé par discipline aussi: la linguistique n'abandonne jamais la quantification, et l'archéologie non plus, avec le maintien du champ de l'archéométrie, qui a sa *Revue d'archéométrie* en 1977 (Bourin et

Zadora-Rio, 2013).

Surtout, les années 1980 et 1990 sont une période de développements en ce qui concerne la représentation numérique des sources et données historiques, leur encodage et leur échange, avec l'émergence successive, d'une part, du modèle relationnel (*BASE DE DONNÉES), et de l'autre, pour les données textuelles, des métalangages SGML, puis XML, ainsi que de la *Text Encoding initiative* (*TEXTE NUMÉRIQUE). C'est aussi la période du développement de l'informatisation des institutions patrimoniales, à l'instar de la BnF dont le catalogue informatisé, soutenu justement par E. Le Roy Ladurie, est inauguré en 1988 (Lemny, 2017).

Le rôle joué par le développement de la micro-informatique personnelle est plus ambivalent: par la démocratisation de l'accès, il a pu encourager l'innovation et la démultiplication des expérimentations, mais il vient aussi perturber l'écosystème des centres de calcul, lieux de sociabilité et de partage d'expertise (Hockey, 2004), et entraîne des résistances contre cette tendance décentralisatrice dans des institutions comme le CNRS (Genet, 2003). Surtout, c'est le développement du web qui paraît être décisif dans un renouveau du numérique en histoire, tout en entraînant une focalisation sur la publication en ligne (Hockey, 2004) et la diffusion tant d'éditions de sources que d'instruments de recherche. Sur le site de l'École des chartes, par exemple, les premiers dossiers documentaires en ligne apparaissent à partir de 2001, bientôt suivis par les plateformes Thélème et ÉLEC de matériaux pédagogiques et d'éditions électroniques, et notamment le projet d'édition du Cartulaire blanc de l'abbaye de Saint-Denis (Guyotjeannin et Poupeau, 2003).

La révolution portée par le web se révèle ainsi être surtout celle de la fluidité, voire du « déluge de données », qui va permettre l'émergence d'un nouveau paradigme de recherche, théorisé par Jim Gray sous le nom d'« eScience », de « Data intensive science » ou, plus simplement, en français, de science des données (Grey, 2009; Bell, Hey et Szalay, 2009). Celle-ci unit méthodes statistiques, intelligence artificielle, masse de données disponibles et expertise thématique du chercheur dans une nouvelle méthodologie, applicable tant aux sciences sociales qu'à celles du vivant, qui part des données pour arriver à des énoncés scientifiques, à rebours d'une démarche fondée sur le test d'hypothèses et le primat de la théorie. Au-delà de la naïveté, voire du retour à l'inductionnisme que pourrait opérer cette approche tirée jusqu'à l'excès, la théorisation de Jim Gray rend bien compte de la position absolument centrale de la question des données, du renouveau méthodologique qui l'entoure, et des bénéfices possibles tant en termes de reproductibilité ou de cumulabilité, que tout simplement de l'émergence de connaissances nouvelles fondées sur l'appréhension de séries très importantes. Ce renouveau de l'approche quantitative et sérielle, devenue matériellement et techniquement plus aisément implémentable, trouve des échos théoriques, par exemple dans la remise en cause des études littéraires fondées sur un tout petit canon de textes, plutôt qu'une appréhension globale de la production d'une époque donnée, chez un chercheur comme Moretti (Moretti, 2013), et elle entretient également des

espoirs de gigantisme tels que ceux des projets de *Venice Time Machine* (Kaplan, 2015), voire d'*European Time Machine*.

Plus récemment, on observe également un renouveau de l'intérêt pour l'étude des systèmes complexes, la modélisation mathématique et la simulation, qui, venus de la physique statistique, trouvent également des applications emblématiques dans les sciences du climat ou l'épidémiologie, tout en se développant aussi en économie et dans les sciences sociales, et, de là, les sciences historiques. Assez embryonnaires dans ce domaine, elles fournissent des perspectives intéressantes pour passer d'approches descriptives des faits individuels à une démarche explicative cherchant à dégager des mécanismes ou des structures récurrentes en en fournissant une interprétation (Barceló et Del Castillo, 2016).

Jean-Baptiste Camps

Références

En pied de notice

Bell (Gordon), Hey (Tony), et Szalay (Alex), « Beyond the Data Deluge », *Science*, 323-5919 (2009), p. 1297-1298.

Bourin (Monique) et Zadora-Rio (Élisabeth), « La mesure au Moyen Âge et les mesures des médiévistes: remarques en forme de conclusion », dans *Mesure et histoire médiévale: XLIIIe Congrès de la SHMESP (Tours, 31 mai-2 juin 2012)*, Paris, 2013, p. 374-396.

Chaunu (Pierre), « Les dépassements de l'histoire quantitative: retrospective et perspective », *Mélanges de la Casa de Velázquez*, 8 (1972), p. 647-685, doi: 10.3406/casa.1972.1072.

Cinquin (Sophie), « Utiliser la lexicométrie en histoire (1): panorama historiographique », *Devenir historien-ne*, 2011, <https://devhist.hypotheses.org/898> (consulté le 14 mars 2020).

Furet (François), « Histoire quantitative et construction du fait historique », *Annales: économies, sociétés, civilisations*, 26-1 (1971), p. 63-75, doi: 10.3406/ahess.1971.422459.

Lemercier (Claire) et Zalc (Claire), *Méthodes quantitatives pour l'historien*, Paris, 2007 (Repères, 507).

Moretti (Franco), *Distant reading*, Londres, 2013.

Schreibman (Susan), Siemens (Raymond Georges) et Unsworth (John), *A Companion to Digital Humanities*, Oxford, 2004, <http://www.digitalhumanities.org/companion>.

—, *A New Companion to Digital Humanities*, Malden, 2016.

En bibliographie générale

Barceló (Juan A.) et Del Castillo (Florencia), *Simulating Prehistoric and Ancient Worlds*, 2016.

Bell (Gordon), Hey (Tony), et Szalay (Alex), « Beyond the Data Deluge », *Science*, 323-5919 (2009), p. 1297-1298.

Bourin (Monique) et Zadora-Rio (Élisabeth), « La mesure au Moyen Âge et les mesures des médiévistes: remarques en forme de conclusion », dans *Mesure et histoire médiévale: XLIIIe Congrès de la SHMESP (Tours, 31 mai-2 juin 2012)*, Paris, 2013, p. 374-396.

Bullough (V.), « The Computer and the Historian: Some Tentative Beginnings », *Computers and the Humanities*, 1-3 (1967), p. 61-64, www.jstor.org/stable/30199211.

Chaunu (Pierre), « Les dépassements de l'histoire quantitative: retrospective et perspective », *Mélanges de la Casa de Velázquez*, 8 (1972), p. 647-685, doi: 10.3406/casa.1972.1072.

Cinquin (Sophie), « Utiliser la lexicométrie en histoire (1): panorama historiographique », *Devenir historien-ne*, 2011, <https://devhist.hypotheses.org/898> (consulté le 14 mars 2020).

Couturier (Marcel), « Vers une nouvelle méthodologie mécanographique: la préparation des données », *Annales: économies, sociétés, civilisations*, 21-4 (1966), p. 769-778, doi: 10.3406/ahess.1966.421421.

Crevier (Daniel), *AI: The Tumultuous Search for Artificial Intelligence*, New York, 1993.

Daumard (Adeline) et Furet (François), « Méthodes de l'Histoire sociale: les Archives notariales et la Mécanographie », *Annales: économies, sociétés, civilisations*, 14-4 (1959), p. 676-693, doi: <https://doi.org/10.3406/ahess.1959.2865>.

Dearing (V.A.), *Introduction to Computer Programming for Literary Students*, Berkeley, 1967.

Delcourt (Christian), « 1984 en Belgique », *Revue Informatique et Statistique dans les Sciences humaines*, 1984, p. 223-229, <http://web.philo.ulg.ac.be/rissh/wp-content/uploads/sites/10/pdf/Annee1984/Notes/CDelcourt.pdf>.

Delcourt (Christian), « Stylometry », *Revue belge de philologie et d'histoire*, 80 (2002), p. 979-1002, DOI: 10.3406/rbph.2002.4651.

Delumeau (Jean), « Méthode mécanographique et trafic maritime: les terre-neuvers malouins à la fin du XVIIe siècle », *Annales: Histoire, Sciences Sociales*, 16-4 (1961), p. 665-685.

EADH, « Wilhelm Ott », s.d., <http://eadh.org/wilhelm-ott>, consulté le 14 mars 2020.

Froger (Jacques), *La Critique des textes et son automatisation*, Paris, 1968 (Initiation Aux Nouveautés de La Science).

- Furet (François), « Histoire quantitative et construction du fait historique », *Annales: économies, sociétés, civilisations*, 26-1 (1971), p. 63-75, doi: 10.3406/ahess.1971.422459.
- Gardin (J.-C.) & Garelli (P.), « Étude des établissements assyriens en Cappadoce par ordinateur », *Annales*, 16-5 (1961), p. 837–876.
- Genet (Jean-Philippe), « Histoire, Informatique, Mesure », *Histoire & Mesure*, 1-1 (1986), p. 7-18, DOI : <https://doi.org/10.3406/hism.1986.904>.
- Genet (Jean-Philippe), *La Génèse de l'Etat moderne*, propos recueillis par Peter Stockinger, 2003, en ligne: <http://www.archivesaudiovisuelles.fr/123/>.
- Gray (Jim), « Jim Gray on eScience: A transformed scientific method », dans *The fourth paradigm: Data-intensive scientific discovery*, T. Hey, S. Tansley, K. Tolle (éd.), Washington, 2009.
- Greg (Walter Wilson), *The Calculus of Variants: An Essay on Textual Criticism*, Oxford, 1927.
- Guyotjeannin (Olivier) et Poupeau (Gautier), « Le projet d'édition électronique du cartulaire blanc de l'abbaye de Saint-Denis et les projets électroniques de l'École nationale des Chartres », *Le médiéviste et l'ordinateur*, 42 (2003), num. spécial *La diplomatie*, p. 93-99.
- Heller (J.) et Logemann (G.W.), *Computer Methods in the Humanities*, New York, 1967.
- Hockey (Susan), « The history of humanities computing », *A companion to digital humanities...*, 2004, p. 3-19.
- Hutchins (W. John), « Machine Translation over fifty years », *Histoire Épistémologie Langage*, 23-1 (2001), p. 7-31, doi: 10.3406/hel.2001.2815.
- Kaplan (Frédéric), « The Venice Time Machine », *Proceedings of the 2015 ACM Symposium on Document Engineering*, 2015, p. 73-73.
- The Computer and Literary Style*, ed. Jacob Leed, Kent State Univ. Press, 1966.
- Lemny (Stefan), « 'L'historien de demain sera programmeur...': Emmanuel Le Roy Ladurie et les défis de la science », *L'Histoire à la BnF*, 2017, <https://histoirebnf.hypotheses.org/1505> (consulté le 24/02/2020).
- Léon (Jacqueline), « Le traitement automatique des langues: présentation », *Histoire Épistémologie Langage*, 23-1 (2001), p. 3-6, doi: 10.3406/hel.2001.2814.
- Le Roy Ladurie (Emmanuel), « La fin des érudits », *Le Nouvel Observateur*, 8 mai 1968; republ. sous le titre « L'historien et l'ordinateur », dans *Le territoire de l'historien*, Paris, 1973, p. 14-...
- Maas (Paul), « Leitfehler Und Stemmatische Typen », *Byzantinische Zeitschrift*, 37-2 (1937), p. 289–94, doi:10.1515/byzs.1937.37.2.289.

Mendenhall (T. C.), « A Mechanical Solution of a Literary Problem », *The Popular Science Monthly*, 60 (1901), p. 97-105.

Moretti (Franco), *Distant reading*, Londres, 2013.

Muller (Charles), *Initiation à la statistique linguistique, ...*, 1968

Milic (Louis T.), «The Next Step», *Computers and the Humanities*, 1-1 (1966), p. 3-6, <http://www.jstor.org/stable/30199191>.

Prost (Antoine), *Vocabulaire des proclamations électorales de 1881, 1885 et 1889*, Paris, 1974.

Robert (Martin), « Hommage à Paul Imbs (1908-1987) », *Comptes rendus des séances de l'Académie des Inscriptions et Belles-Lettres*, 154-1 (2010), p. 42-47, doi: 10.3406/crai.2010.92772.

Robin (Régine), *Histoire et linguistique*, Paris, 1973.

Shorter (Edward), *The Historian and the Computer: a practical Guide*, Englewood Cliffs, 1971.

Tannery (Paul), « La Stylométrie : ses origines et son présent », *Revue Philosophique de la France et de l'Étranger*, 47 (1899), p. 159-169, <http://www.jstor.org/stable/41076911>.

Quentin (Henri), *Essais de critique textuelle (ecdotique)*, Paris, 1926.