



21
juin
2019

*Journée d'études
du groupe régional Aquitaine
De l'Association des archivistes français
Organisée en partenariat
avec l'axe E3D du MICA*

Le document, parent pauvre de la donnée L'archive à l'âge du numérique

- ▶ Université Bordeaux Montaigne
Maison des sciences de l'Homme d'Aquitaine
- ▶ Organisation et contact :
Franck Cormerais - franck.cormerais1@gmail.com



« *Il n'y aura de science sociale, à mon sens, que dans une réconciliation, une pratique simultanée de nos divers métiers. Les dresser l'un contre l'autre, chose facile, mais cette dispute se joue sur de bien vieux airs. C'est d'une musique nouvelle que nous avons besoin* »

Fernand BRAUDEL, chapitre IV de l'Introduction du *Traité de sociologie*, publié sous la direction de Georges Gurvitch, Paris, PUF, 2 vol., 516 et 466 p., in-8° ; 1^{ère} édition 1958-1960, 3^e édition, 1967-1968 [repris dans : *Ecrits sur l'Histoire*, Paris, Champs Flammarion, 1969, p. 120-121].

Interrogeons-nous, collectivement, sur les jeux d'échelle possibles dans l'abord présent des fonds d'archives et des documents disponibles, compte tenu de deux phénomènes distincts : une informatisation des inventaires qui a commencé en France dans les années 1980 et une mise en ligne des images d'une partie des originaux eux-mêmes dans les fonds qui a, dès les années 1990, représenté une autre évolution majeure. Cette évolution impressionnante, qui conduit à évoquer un « âge numérique », ne recouvrait pas, et de loin, la totalité d'un matériau accumulé depuis plus de deux siècles dans le réseau national des archives, avec une continuité permise par la forme des institutions françaises et l'extraordinaire densité de la production, dans ce pays, des archives.

De même que dans l'environnement écrit de la production des archives - institutionnel, administratif, judiciaire ou privé - une première transformation / sélection est opérée par les producteurs eux-mêmes de cette masse immense de documents, l'opération de leur collecte et de leur classement pour mise à disposition introduit elle aussi de profondes modifications. Et toutes ces modifications affectent un matériau qui demeure lui-même, aujourd'hui, une partie immergée dans les dépôts, atteignable seulement en salle de lecture du fait que la numérisation n'est pas à l'échelle des masses conservées (et toujours en expansion), de l'« iceberg » des fonds et documents d'archives effectivement disponibles (on line et off line) du Moyen Age à nos jours.

L'opération historiographique à son tour, comme la plus modeste des recherches d'érudition locale ou de généalogie personnelle se constitue elle-même par un choix dans cet ensemble, un déplacement, une série de découvertes d'éléments neufs au milieu de l'ensemble, puis une restitution philologique du matériau à chaque fois sélectionné et surtout une critique. Celle-ci constitue le soubassement d'un récit toujours renouvelé où se mêlent depuis longtemps d'autres documents que ceux des archives (le paysage et son évolution, les faits de langue, les restes archéologiques ...).

L'empirisme et le positivisme avaient rêvé dès le XIX^e siècle d'une mise en oeuvre exhaustive des documents, porteurs d'un paysage historique simplement observé puis finalement complètement décrit, un jour ou l'autre, par des savants capables de s'effacer eux-mêmes devant le « résultat » de leur travail. Cette vision qui fut celle de Langlois et Seignobos dans leur Introduction aux études historiques (1898) évoque aujourd'hui l'acceptation empirique de « résultats » de requêtes dans les bases de

données, d'une recherche qui ne trouverait que ce qu'elle a tout d'abord cherché alors qu'elle ne fait très fréquemment, en archives, que chercher le sens de ce qu'elle a trouvé, au long des cartons et des registres.

Nous savons bien aujourd'hui, en outre, que l'effort critique n'est vraiment fécond qu'à condition qu'il accepte le caractère nécessairement lacunaire de tout corpus obtenu par la recherche au milieu de ce qui est disponible, dans les fonds mêmes et de manière nécessairement plus restreinte, dans les images d'archives en ligne.

Pour se mouvoir des ensembles qu'il aperçoit aux détails des documents où il replonge sans cesse, détails qui parfois ne s'annexent pas mais dérangent, comme l'écrit Arlette Farge, parce qu'ils représentent les écarts et la singularité de chaque vie humaine), le chercheur d'aujourd'hui manque peut-être souvent, en ligne, d'une vue plus claire de l'échelle du fonds. Cette réalité familière à l'archiviste est à la fois floue dans ses contours et vivace dans sa matérialité. Jamais totalement en ligne par définition puisque souvent encore « ouvert », susceptible de compléments à collecter, qui parfois l'expliquent, ou « virtuel » lui-même bien que tout à fait matériel, par défaut de classement et d'identification, le fonds pose comme le document lui-même des problèmes à la fois théoriques (sa *définition*) et pratiques.

Son traitement s'inscrit dans une historicité qui redouble celle de la production elle-même des documents qu'il comporte, et celle des multiples aventures qu'il traverse dans le temps long, depuis sa production, jusqu'à son éventuel dépouillement.

En ce sens, l'abord du fonds et celui du document ne sont jamais complets avec les seules ressources d'une indexation au sein d'un web sémantique qui parfois font écran à leur matérialité et à sa perception par le public.

Programme de la journée

• 09h00 : Ouverture

Franck CORMERAIS, professeur de Sciences de l'information et de la communication à l'Université Bordeaux-Montaigne
Georges CUER, conservateur d'archives (Archives départementales de la Gironde)

• 09h30 - 10h00

Jonathan BARBIER, docteur en histoire contemporaine (2016)
et Antoine MANDRET-DEGEILH, docteur en science politique de Sciences Po Paris (2015)
(Laboratoire des sciences sociales du politique ; Toulouse)
Le vocabulaire des archives à l'épreuve de leur double vie analogique et digitale : ruses et manières propres de cheminer à travers la forêt des produits proposés

• 10h00 - 10h30

Elsa TADIER, docteur en sciences de l'information et de la communication (GRIPIC, Groupe de recherches interdisciplinaires sur les processus d'information et de communication à l'Ecole des hautes études en sciences de l'information et de la communication - CELSA, Paris-Sorbonne)
Mots du corps, mots du livre, mots des archives : enjeux de la transformation des gestes savants ordinaires « à l'heure du numérique »

• 10h30 - 10h45 : Pause

• 10h45 - 11h15

Franck CORMERAIS, professeur à l'Université Bordeaux-Montaigne,
et Lucie VIEILLECROZE, doctorante en Sciences de l'information et de la communication (Laboratoire MICA, Médiation, information, communication, arts ; Bordeaux)
Le triptyque documents, archives, données : une approche du nexialisme à travers le classement du fonds Robert Escarpit

• 11h15 - 11h45

Philippe RYGIEL, Professeur d'histoire contemporaine à l'Ecole normale supérieure de Lyon (LARHA, laboratoire de recherche historique Rhône-Alpes, pôle Histoire numérique)
Les portails Francearchives et Online Archives of California et les deux visions de la normalisation archivistique : mon oncle d'Amérique ?

• 11h45 - 12h15

Yann POTIN, chargé d'études documentaires aux Archives nationales, maître de conférences associé en histoire du droit à l'Université Paris-Nord
Le projet Open Jérusalem Archives pour une nouvelle histoire de la ville : ouvrir, entrouvrir, connecter, communiquer les fonds d'archives

• 12h30 - 14h00

Possibilité de déjeuner au restaurant **Club House** du BEC
Avenue Recteur Jean Babin à Pessac (derrière la MSHA).
Menu du jour Entrée + Plat + Dessert = 13,50 euros à régler sur place
Lien vers le site : <http://www.bec-bordeaux.fr/club-house/restaurant-2/>

• 14h00 - 14h30

Michelle BUBENICEK, directrice de l'Ecole nationale des chartes
et Elsa MARGUIN-HAMON, directeur de la recherche et des relations internationales (Ecole nationale des chartes)
De Theleme à Adele : évolution des plateformes de mise en ligne des supports d'enseignement de l'Ecole nationale des chartes

• 14h30 - 15h00

Fabrice PAPY, professeur à l'Université de Lorraine, Sciences de l'Information et de la communication
Les bases de connaissances en bibliothèque : un modèle innovant pour les services d'archives et leurs fonds, « peuplés d'individus »

• 15h00 - 15h15 : Pause

• 15h15 - 15h45

Jean-Marc WELLER, chercheur CNRS (LISIS : Laboratoire Interdisciplinaire Sciences Innovations Société, UPEM : Université Paris Est Marne)
Ce que la numérisation des écrits fait au travail. Un regard sociologique sur les métiers du droit et leur production d'actes

• 16h15 - 16h45

Victor LE BRETON-BLON, doctorant en Histoire du droit (Institut de recherche Montesquieu, Université de Bordeaux)
L'autre côté du miroir. Du document à la donnée et de la donnée au document, un témoignage d'histoire du droit des affaires à l'époque moderne

Résumés des communications

du 21 juin 2019

- Jonathan BARBIER, docteur en histoire contemporaine (2016) et Antoine MANDRET-DEGEILH, docteur en science politique de Sciences Po Paris (2015) - Laboratoire des sciences sociales du politique, Toulouse.

Le vocabulaire des archives à l'épreuve de leur double vie analogique et digitale : ruses et manières propres de cheminer à travers la forêt des produits proposés

Cette communication vise à interroger la pluralité des mots relatifs au travail sur archives et de leurs acceptions. La numérisation, la photographie, les délais de plus en plus courts des séjours aux archives ont changé en effet en profondeur le lien, le contact, voire le toucher que tout lecteur entretient avec les documents d'archives. Le vocabulaire se fait le reflet de ces changements : il suffit ici de penser à des verbes tels que « communiquer », « publier », « numériser », « dépouiller », dont l'usage immodéré - et souvent irréfléchi - se développe dans le langage courant, pour s'en convaincre. Au moment où la « révolution » numérique bouleverse la pratique du travail sur archives, il apparaît donc opportun de s'interroger sur les lexiques des archives que lecteurs et archivistes emploient. Le « fonds » a-t-il encore un sens à l'ère de la numérisation massive des pièces ? Les « documents » sont-ils toujours pertinents pour qualifier les archives ? Les « données » sont-elles nécessairement des archives ? etc. Sans chercher à l'exhaustivité, nous désirons appréhender cette problématique des mots et de leurs définitions en nous intéressant à la fois aux textes et aux gestes des praticiens.

Pour ce faire, nous comparerons les usages de termes précédemment mentionnés par les différents praticiens du monde des archives. Il s'agira aussi bien de revenir sur les évolutions lexicologiques et sémantiques des textes législatifs qui accompagnent les transformations de l'archivistique - en allant voir notamment du côté des coulisses de leur élaboration -. que d'analyser les changements qui ont affecté ces dernières décennies les lexiques des spécialistes des archives, qu'il s'agisse des archivistes eux-mêmes ou bien des lecteurs des archives. Ces derniers ont en effet eux aussi appris à forger leurs propres boîtes à outils lexicologiques, lesquelles ne manquent pas de varier selon les publics de lecteurs et les usages que ces publics font des archives. Une généalogiste ne parle pas, en effet, de la même façon du travail sur archives qu'un.e chercheur.e et ce, d'autant plus que les chercheur.e.s jonglent avec différents lexiques selon leur discipline de rattachement notamment. Historiens, sociologues, politistes ont en effet des conceptions différentes des archives en fonction de leurs terrains d'enquête, de leurs méthodologies ou de leurs formations respectives.

- Elsa TADIER, docteur en sciences de l'information et de la communication (GRIPIC, Groupe de recherches interdisciplinaires sur les processus d'information et de communication à l'École des hautes études en sciences de l'information et de la communication - CELSA, Paris-Sorbonne)

Mots du corps, mots du livre, mots des archives : enjeux de la transformation des gestes savants ordinaires «à l'heure du numérique »

La pratique des archives nous invite à plonger dans une langue, un vocabulaire de spécialistes qui partage avec celle du livre une forme d'intimité, sinon à tous niveaux, du moins à plusieurs égards.

Pour le livre, ce vocabulaire (depuis la codicologie jusqu'aux acteurs de sa fabrication) relève fortement d'une corporéité endormie dans une langue qui l'éveille, en réveille les infimes vibrations. Le livre est en effet considéré comme un *corps* pourvu d'une *tête* et de *pieds*, d'un *dos* et d'une *épine*, notamment.

Au-delà de la projection anthropomorphique, la langue du livre - comme parfois celle des archives - *nous invite à revenir à la matérialité des objets que nous manipulons* dans la pratique de recherche : par exemple, la *peau* du parchemin, caractérisée par ses *nervures* et *cicatrices*, sont autant d'éléments dont le lecteur - le chercheur ou l'archiviste qui classe, organise, recense, thématise, traite le document - se font les témoins privilégiés. Cette matérialité singulière n'est pas dénuée de sens : elle travaille notre appréhension, notre compréhension des documents constitués pour nous en archives. Elle en donne des indices : temporalités, accidents, ratures, effacements, métamorphoses s'y lient et s'y délient...

Ainsi la langue convoque-t-elle le corps du lecteur qui, lui aussi, « lit », en mobilisant dans sa lecture quantité d'éléments relevant d'un savoir « infra-ordinaire » : il n'est que de penser au papier qui se touche, se sent, se définit à travers la *main*, dont on a fait un geste de mesure pour évaluer son poids et son grammage, mesure qui a aujourd'hui valeur arithmétique.

Après un retour sur cette « langue du corps », l'intervention visera à soulever un certain nombre d'interrogations autour des métamorphoses engagées par le numérique. Depuis que s'est amorcé un vaste mouvement de numérisation de certains documents d'archives dont l'image devient « consultable » de n'importe où, certaines archives (et non les fonds) deviennent « accessibles » partout où elles ne sont pas. Le corps du lecteur n'a plus besoin de se déplacer pour les faire venir à lui. Mais fait-il vraiment advenir les archives au sens plein du terme ? Que se joue-t-il dans cette transformation ? La numérisation engage en effet une *métamorphose matérielle* qui transforme le rapport à la pratique des archives - et à la production du savoir. Au cœur de cette problématique : la transformation du *corps* à *corps* des archives et de leur lecteur. Le corps sera-t-il la langue perdue du numérique ?

- Franck CORMERAIS, professeur à l'Université Bordeaux-Montaigne, et Lucie VIEILLECROZE, doctorante en Sciences de l'information et de la communication (Laboratoire MICA, Médiation, information, communication, arts ; Bordeaux-Montaigne)

Le triptyque documents, archives, données : une approche du nexialisme à travers le classement du fonds Robert Escarpit

Dans le texte *Pour une épistémologie des sciences de l'information et de la communication* (1976), Robert Escarpit assimile les sciences de l'information et de la communication au nexialisme, une science fictive imaginée par A. E. Van Vogt dans son livre *La Faune de l'espace*, dont l'objet est de faire le lien entre les différents savoirs. Cette proposition fait écho à l'activité du chercheur qui s'est lui-même intéressé à différents champs scientifiques émergents (littérature comparée, sociologie de la littérature, sciences de l'information et de la communication, etc.). La diversité de ses travaux transparaît aujourd'hui dans le fonds d'archives Robert Escarpit et le respect de l'intégrité du fonds implique la cohabitation de ces différentes disciplines au sein des documents.

En prenant l'exemple du fonds d'archives Robert Escarpit, il s'agit de s'interroger sur les apports d'une classification archivistique pour la modélisation de l'activité scientifique d'un chercheur et sur les apports et limites du numérique pour le classement et la mise en données de ses archives. Alors que le classement est ce qui transforme la pile de documents en archives et permet la mise en données du fonds, comment les technologies numériques peuvent-elles favoriser la recherche et la consultation des archives ?

- Philippe RYGIEL, Professeur d'histoire contemporaine à l'École normale supérieure de Lyon (LARHA, laboratoire de recherche historique Rhône-Alpes, pôle Histoire numérique)

Les portails Francearchives et Online Archives of California et les deux visions de la normalisation archivistique : mon oncle d'Amérique ?

- Yann POTIN, chargé d'études documentaires aux Archives nationales, maître de conférences associé en histoire du droit à l'Université Paris-Nord

Le projet *Open Jérusalem Archives* pour une nouvelle histoire de la ville (sainte) : ouvrir, retrouver, connecter, communiquer les fonds d'archives

Cette intervention entend constituer un retour réflexif et méthodologique autour d'une enquête collective internationale, associant une équipe d'une dizaine de chercheurs et archivistes, mobilisant un réseau d'une cinquantaine de collaborateurs. Le point de départ est une question historiographique classique - quelles sont les documents d'archives mobilisables pour faire l'histoire locale d'une ville globale comme le fut Jérusalem au XIXe siècle ? C'est à ce défi qu'a voulu répondre le programme « Open Jerusalem Archives (1840-1940) », financé par l'ERC (European Research Council) entre 2014 et 2018 et porté par Vincent Lemire (voir le carnet de recherche : <https://openjlem.hypotheses.org/>). Il s'agissait autant de se libérer d'une lecture univoque par fonds « nationaux » que de trouver un moyen de contourner la dispersion géographique de documents disséminés à travers la Turquie, divers pays d'Europe et du Moyen-Orient, mais aussi la Russie, l'Arménie ou l'Éthiopie, tout en prenant cette même logique de distribution pour objet de réflexion et d'investigation.

Au total, le produit principal de la recherche, à partir d'une série d'opérations d'identification, sinon de collecte, de cartographie et de descriptions archivistiques d'une masse documentaire fondamentalement hybride sur le plan juridique et plus encore multilinguistique, a constitué à la mise en œuvre d'une sorte d'hyper inventaire raisonné des fonds, structuré par une relative uniformité de description et une véritable unité de description linguistique (<http://www.openjerusalem.org/>).

En reconstituant de manière rétroactive les grandes étapes de réalisation du projet, on s'efforcera de restituer la confrontation et les malentendus entre producteurs et usagers, entre archivistes et chercheurs, entre mise en archives et fabrication des données.

- Michelle BUBENICEK, directrice de l'École nationale des chartes et Elsa MARGUIN-HAMON, directrice des études (École nationale des chartes)

De Theleme à Adele : évolution des plateformes de mise en ligne des supports d'enseignement de l'École nationale des chartes (titre provisoire)

L'École nationale des chartes s'est engagée de longue date dans la numérisation et la mise en ligne de ses supports d'enseignement traditionnels que sont les fac similés. Dès les années 2000 le site THELEME (Techniques pour l'Historien en Ligne: Etudes, Manuels, Exercices, Bibliographies) permet la consultation de ces documents dans la perspective pédagogique assumée d'en améliorer la compréhension, à tous niveaux: paléographique, linguistique, diplomatique. Outre les dossiers de documents, transcrits et commentés au crible de différentes disciplines et méthodes critiques, le site propose des cours d'introduction et des outils pour l'étude et la recherche permettant de situer les objets textuels dans le contexte épistémologique qui est le leur.

L'évolution des techniques et les progrès de l'intelligence artificielle dans le domaine de la reconnaissance des écritures, de la lemmatisation des textes et du traitement automatisé de la langue permet à ces ressources pédagogiques de faire peau neuve dans le cadre du programme ADELE (Album de Diplomatique en ligne), que l'ENC vient de déployer dans le cadre des cours de diplomatique d'Olivier Guyotjeannin.

L'intervention reviendra sur les perspectives qu'ouvre très largement cette nouvelle plateforme.

- Fabrice PAPY, professeur à l'Université de Lorraine, Sciences de l'Information et de la communication

Les innovations numériques anthropocentrées pour le web des données et des documents : davantage de liens, davantage de sens pour les usagers ?

Les développements récents en matière de dispositifs numériques technodocumentaires révèlent que les approches technocentrées ne peuvent plus être l'alpha et l'omega de ces réalisations numériques destinées à un très large public dont l'empan des affiliations et des habiletés techniques/technologiques s'avère fort différent. La pénétration massive des réseaux sociaux numériques dans les activités les plus diverses et la généralisation de l'Expérience Utilisateur (UX) ont déplacé l'épicentre de conception d'une approche technocentrée à une approche résolument anthropocentrée où l'activité du sujet, ses capacités et son projet d'usage ne peuvent plus être écartées ou absorbées dans des conceptions modélisatrices excluant usages et détournements. Le frémissement d'une réappropriation des objets numériques et de leurs traitements par des usagers invités à en imaginer des usages inédits (voire innovants) trouvent voix au chapitre, à l'instar des propositions citoyennes de la plate-forme data.gouv.fr. Les recherches en Sciences de l'Information et de la Communication interrogent depuis plusieurs années ces processus d'information et de communication instrumentés par le numérique où les questions d'usage sont prépondérantes. Je m'attacherai au travers de quelques expérimentations dans différents contextes documentaires récents à illustrer les approches innovantes destinées à améliorer les usages de ces plates-formes numériques de notre « société de l'information ».

- Jean-Marc WELLER, chercheur CNRS (LISIS : Laboratoire Interdisciplinaire Sciences Innovations Société, UPEM : Université Paris Est Marne)

Ce que la numérisation des écrits fait au travail. Un regard sociologique sur les métiers du droit et leur production d'actes

Très étroitement associé à l'écrit et à la preuve, le document fait l'objet d'un travail de confection, d'enregistrement et d'interprétation éminemment déterminant. Si les juristes, les médecins, les bibliothécaires, les chercheurs, et bien d'autres professionnels, n'ont de cesse d'en manipuler les écrits, ils le font de manière différente, à des fins qui leur sont propres, et au prix de procédés et de conceptions

qui leur confèrent des statuts différents. Ces différences apparaissent particulièrement manifestes dès lors qu'on s'intéresse à la matière dont les écrits sont fait et à ce celle-ci fait à la connaissance.

Dans le prolongement des travaux sociologiques qui se sont intéressés aux écrits que je rappellerai brièvement, dans la perspective d'un dialogue avec la question des archives et de leur numérisation posée par la journée d'étude, mon intervention reposera sur le cas des métiers juridiques à l'épreuve de la digitalisation. J'examinerai, en particulier, l'intérêt et les limites d'une analyse du travail opposant voie analogique et voie numérique dans le traitement des écrits.

- Victor LE BRETON-BLON, doctorant en Histoire du droit (Institut de recherche Montesquieu, Université de Bordeaux)

L'autre côté du miroir : une thèse d'histoire du droit des affaires à l'époque moderne, du document à la donnée et de la donnée au document.

Le quotidien de l'historien renouvelle sans cesse la question de la diffusion et de la médiation des sources historiques. Cependant, ces thématiques constamment réitérées au fil des progrès techniques peuvent rapidement se heurter à la complexité d'un fonds d'archives.

L'exploitation des archives est en grande partie tributaire de deux notions : le document et la donnée. Le juriste historien collationne ainsi les documents et les systématise pour en extraire une expression du droit. La perspective d'une numérisation des archives ou d'un classement axiomatique et informatisé de leurs ressources peut alors apparaître comme une aubaine pour le chercheur pressé. Toutefois, la crainte d'un document perdu dans sa donnée se profile. La disparition du matériel physique au profit d'un chiffre, d'un pourcentage, pourrait être la résultante d'une perte d'appréhension matérielle de l'acte qui est pourtant nécessaire à l'historicisation du document.

Ainsi, les différentes orientations qui pourront naître des discussions suscitées par ce colloque ne doivent pas masquer la dialectique du document et de la donnée. Le document n'est-il qu'un acte qui peut être transcrit et détaché de son support physique ? La donnée est-elle un instrument suffisant pour le chercheur ? A-t-elle gagné son autonomie grâce aux lettres de noblesse de l'analyse quantitative ?

Notre communication ambitionne donc de replacer au centre du débat lié à la numérisation des archives la question du document, de la donnée et des liens qu'entretiennent ces deux éléments à travers le témoignage d'une thèse d'histoire du droit.



► Organisation et contact :
Franck Cormerais - franck.cormerais1@gmail.com

