QUESTIONNER LES HUMANITÉS NUMÉRIQUES

Sous la direction scientifique de
Françoise PAQUIENSÉGUY et Nicolas PÉLISSIER

SFSIC
SOCIÉTÉ FRANÇAISE DES SCIENCES DE L'INFORMATION ET DE LA COMMUNICATION

Conférence permanente des directeurs·trices de laboratoires en sciences de l'information et de la communication
CPDirSIC

QUESTIONNER LES HUMANITÉS NUMÉRIQUES

Positions et propositions des Sic

Création et mise en page

Atelier Congard
www.atelier-congard.fr

© SFSIC et CPdirsic - 2021.

ISBN : 978-2-914872-00-3

Sous la coordination de
Nicolas Pélissier et Françoise Paquienséguy

QUESTIONNER LES HUMANITÉS NUMÉRIQUES

Positions et propositions des Sic

SFSIC
SOCIÉTÉ FRANÇAISE DES SCIENCES DE L'INFORMATION ET DE LA COMMUNICATION

Conférence permanente des directeurs-trices de laboratoires en Sciences de l'information et de la communication
CPDirSIC

Sommaire

Avant-propos **7**
Philippe Bonfils et Jacques Walter

Introduction **9**
Françoise Paquienséguy et Nicolas Pélissier

— PARTIE I —

Vers un renouvellement théorique et épistémique ?

Mise en perspectives **21**
Françoise Paquienséguy

CHAPITRE I | Des enjeux épistémiques aux questions sociopolitiques **25**
Roger Bautier

CHAPITRE II | Les Humanités Numériques, quelles définitions ? **41**
Dominique Carré, Fanny Georges et Jérôme Valluy

CHAPITRE III | Les Humanités Numériques, depuis les Sic **57**
Chérifa Boukacem et Françoise Paquienséguy

— PARTIE II —

Humanités Numériques : de l'éthique au politique

Mise en perspectives **75**
Christine Barats et Ghislaine Chabert

CHAPITRE IV | Situer et contextualiser les discours politiques **77**
Christine Chevret-Castellani

CHAPITRE V | Questionnement éthique par rapport au numérique **97**
Cécile Dolbeau-Bandin, Jean-Claude Domenget et Carsten Wilhelm

CHAPITRE VI | Inscrire les Humanités Numériques dans une économie de la production scientifique **117**
Vincent Bullich

— PARTIE III —

Humanités Numériques, une question de méthodes ?

Mise en perspectives **135**
Brigitte Juanals et Éric Triquet

CHAPITRE VII | Computation et interprétation : les études digitales une opportunité pour penser une rencontre avec les Sic **139**
Franck Cormerais

CHAPITRE VIII | Apports critiques sur les méthodologies liées aux Humanités Numériques **153**
Christine Barats, Amel Fraisse, Martha Severo, Julia Bonaccorsi et Lise Verlaet

CHAPITRE IX | Regards multidimensionnels sur le numérique **171**
Christine Barats, Amel Fraisse, Pascal Marchand, Angeliki Monnier et Lise Verlaet

CHAPITRE X | Pluralité et antériorité des objets et des méthodes **185**
Julia Bonaccorsi et Brigitte Juanals

— PARTIE IV —

Formation aux Humanités Numériques : quelle place et enjeux pour les Sic ?

Mise en perspectives **201**
Franck Debos et Nicolas Pélissier

CHAPITRE XI | Humanités Numériques au travail (HNT) et la professionnalisation des communicateurs **205**
Elise Maas et Sylvie Alemanno

CHAPITRE XII | Cartographie et enjeux des formations en Humanités Numériques **227**
Elie Allouche et Orélie Desfriches-Doria

CHAPITRE XIII | Les enjeux de la formation de formateurs aux Humanités Numériques **257**
Laurent Collet, Thibaud Hulin et Emilie Rémond

CHAPITRE XIV | Parcours et méthodes de formation aux Humanités Numériques Vers un socle commun de compétences ? **275**
Nicolas Pélissier, Éric Triquet et Arnaud Mercier

— CONCLUSION GÉNÉRALE —

Penser, former, transformer **293**
Nicolas Pélissier et Françoise Paquienséguy

Avant-propos

Cet ouvrage est issu d'une collaboration inédite entre la Société Française des Sciences de l'Information et de la Communication (SFSIC) et la Conférence Permanente des Directeur·trices d'unités de recherche en Sciences de l'Information et de la Communication (CPdirsic).

Son contenu résulte de deux journées d'échanges très riches sur les Humanités Numériques. Il s'agit d'un objet scientifique passionnant, qui fait débat, et sur lequel les Sic ont produit de nombreux travaux. Une trentaine de collègues ont ainsi participé à ce projet et livrent leurs contributions dans une perspective critique.

Ce partenariat entre les deux instances démontre, si besoin était, la capacité à se mobiliser pour participer à une réflexion dont l'intérêt dépasse largement le cadre de la discipline.

Merci à Françoise Paquienséguy et Nicolas Pélissier qui ont assuré la coordination scientifique du projet.

Philippe Bonfils
Président de la SFSIC

et **Jacques Walter**
Président de la CPdirsic

INTRODUCTION GÉNÉRALE

Questionner les Humanités Numériques
Un projet disciplinaire

Cet ouvrage est le fruit d'un processus réflexif et délibératif au long cours développé au sein de la communauté épistémique des chercheur.e.s en sciences de l'information et de la communication (Sic), et plus particulièrement celles et ceux qui étudient de près la transformation numérique de nos sociétés. Nous proposons ici d'évoquer la genèse et le cheminement de ce qu'il convient d'appeler une *intention disciplinaire* visant à construire, déconstruire, voire reconstruire un objet scientifique pour le moins controversé.

Un objet protéiforme et complexe à définir

Le projet actuel porté par les *Digital Humanities* (Humanités Numériques) renvoie par sa terminologie à un projet plus ancien, celui de l'humanisme réformé des XV^e^-XVI^e^ siècles, caractérisé par une démarche intellectuelle, d'ordre métapolitique, visant à développer les sciences et techniques au service de la libre circulation des hommes, des biens et des idées. Concept de la Renaissance ou renaissance d'un concept ? Pour les tenants des Humanités Numériques, il s'agirait effectivement d'achever le projet d'Érasme, Galilée ou Copernic, en faisant en sorte que les technologies digitales renouvellent en profondeur les modes d'apprentissage, de traitement et de partage des savoirs, dans le sens d'une plus grande fluidité, transparence voire démocratisation. Mais aussi d'une plus grande scientificité, en élargissant le domaine des humanités traditionnelles (lettres, langues, arts…) à celui des sciences humaines et sociales, voire des sciences de l'ingénierie, dans une dynamique de mise en relation

généralisée des connaissances chère à l'une des plus célèbres figures de la Renaissance : Léonard de Vinci.

À ce jour, la définition la plus stabilisée des Humanités Numériques (telle qu'énoncée dans le « Manifeste » éponyme de 2010) pourrait être la suivante *: usage des technologies numériques par les humanités classiques (lettres, langues, arts) et les sciences humaines et sociales comme méthodes renouvelées de collecte, traitement et diffusion des données.* Cette conception initiale, suite aux échanges qu'elle a suscités dans la communauté scientifique, a débouché sur une conception élargie à deux autres domaines d'activité : le traitement de grands corpus documentaires au service des organisations (*Big Data*), ainsi que l'archéologie des médias et des dispositifs sociotechniques, qui renvoie à une approche *médiologique* telle que préconisée par Régis Debray. Quant aux *Digital Studies* proposées par Bernard Stiegler, elles renvoient à une acception encore plus large : celle de l'étude de l'impact du numérique sur l'ensemble de notre écosystème sociotechnique. De par leur histoire et leur identité disciplinaire, les Sic semblent de situer davantage dans cette dernière perspective… mais pas seulement, comme le montrent de nombreux travaux récents.

Le numérique, une extension du domaine des humanités ?

Au-delà des discours prophétiques à visée performative, que recouvre aujourd'hui le projet des Humanités Numériques dans les pratiques académiques ? Tout d'abord un domaine de recherches interdisciplinaire, croisant l'informatique et autres sciences de l'ingénierie avec les recherches en arts, lettres, langues et sciences humaines et sociales.

Ensuite, de nouvelles formations et nouveaux diplômes, aux contenus très attrayants pour les étudiants mais aux débouchés encore incertains, par rapport à des filières disciplinaires mieux identifiées à ce jour par les employeurs. Enfin, de nouveaux programmes ambitieux de recherche, à vocation interdisciplinaire, permettant un décloisonnement des savoirs et pratiques scientifiques.

Certes ce projet est riche en promesses de différente nature : transformation des pratiques infocomunicationnelles des chercheurs (sources, collecte, traitement, visualisation et valorisation des données…) ; revitalisation de filières pédagogiques aux effectifs et à l'attractivité en déclin ; nouvelles formes de dialogue interdisciplinaire autour de programmes fédérateurs ; possibilités de renouvellement de

l'écriture et de la diffusion de la recherche à partir des outils développés par la science ouverte et collaborative ; mais aussi opportunités de renforcer la scientificité des méthodes en permettant une meilleure complémentarité entre recherches qualitatives et quantitatives.

Mais ces promesses peuvent aussi s'accompagner de possibles dérives : sous-estimation du contexte social et économique de la production des données ; priorisation des approches quantitatives au détriment d'approches plus qualitatives ; hégémonie accrue de l'informatique sur les sciences humaines avec mise en compétition des chercheurs en fonction de leurs usages du numérique ; illusion scientiste quant au rôle des statistiques dans la compréhension des faits sociaux : ou encore risque de privilégier le traitement des données à leur interprétation…

Pour prévenir de tels risques, les enseignants-chercheurs en sciences de l'information et de la communication peuvent jouer un rôle majeur, comme ils ont pu le faire pour des écosystèmes techniques plus anciens, en se penchant plus particulièrement sur la question du sens attribué aux données produites et traitées dans le cadre des programmes Humanités Numériques. Ce constat a d'ailleurs été le point de départ d'une première discussion disciplinaire qui a débouché sur le présent ouvrage et dont nous allons, dans les paragraphes qui suivent, rappeler les principales étapes.

Aux origines du projet, une journée d'études

Intitulée « *Sic et Digital Studies : divergences ou complémentarités* ? », cette manifestation inédite, organisée par la Commission Recherche de la Société Française des Sciences de l'Information et de la Communication (SFSIC), a interrogé le positionnement des Sic sur la question du « numérique », en considérant également celui d'autres disciplines, nouvelles ou plus anciennes, qui tentent de l'appréhender. Elle a prolongé les débats ouverts par une table-ronde portant un regard critique sur les Humanités Numériques et mise en place dans le cadre des Journées Doctorales de la SFSIC en mai 2015 à Lille.

Le 30 mars 2016, accueillie à Paris au Centre National des Arts et Métiers (laboratoire DICEN-IDF, dirigé par Manuel Zacklad), la journée d'études a ouvert la réflexion sur les approches plus spécifiques du « numérique » qui se sont constituées au travers d'études et de programmes de recherche d'envergure sur l'information, les métadonnées et les transformations affectant différentes catégories d'acteurs. D'une part, les médias, qui ont été parmi les premières

industries culturelles à numériser leurs chaînes de production ; d'autre part, les usagers, aux prises avec les bases de données, les équipements, les terminaux, les contenus numérisés et les réseaux ou médias sociaux ; enfin, les discours et les stratégies de l'ensemble des composantes impliquées dans l'écosystème numérique. De nombreuses expressions langagières ont été questionnées par les participant.e.s, de par les logiques qu'elles portent et les enjeux qu'elles impliquent. Sans pouvoir toutes les citer, ces derniers ont évoqué notamment les « territoires numériques », les « objets connectés », les « *big datas* » et autres « *smart cities* », sans oublier les « *livinglabs* » ou « fablabs » associés à l'essor des nouvelles économies « créatives » et « contributives ». Ce nouveau vocable, le plus souvent issu d'un répertoire anglophone, n'a pas manqué d'attirer l'attention, mais aussi de susciter la perplexité des participant.e.s.

De fait, notre communauté scientifique analyse et étudie depuis maintenant plus de trente ans les manifestations, les représentations, les modèles économiques ou les stratégies liées à la dématérialisation de l'information, à la mise en réseau des contenus et programmes, ou encore aux statuts et comportements des usagers. Pour mieux cerner ce domaine de recherches, on peut évoquer trois types de travaux : les études de cas ou enquêtes de terrain longitudinales, qui montrent les changements à l'œuvre dans différents secteurs des industries ou services, et témoignent de leur généralisation ; les montées en théorisation questionnant les théories et les concepts les plus usuels en Sic en les confrontant à une réalité sociale et économique transformée ; enfin, les propositions de problématisation pensées comme des outils spécifiques à la réalité contemporaine, qu'il s'agisse de modèles ou de notions. Indispensables et complémentaires, ces trois approches montrent bien que l'objet des Sic a été transformé, en profondeur comme en extension, par l'écosystème numérique pris dans sa globalité.

Jusqu'à quel point ce dernier imprègne-t-il donc la discipline des Sic ? De fait, on le retrouve dans les terrains, les méthodologies, les théories et définitions proposées par les chercheur.e.s, leurs pratiques professionnelles, ainsi que dans les nouvelles formations proposées par le champ disciplinaire. Mais que sont donc les Sic aujourd'hui, au regard de ces multiples manifestations de la digitalisation ?

La journée du 30 mars 2016 a eu pour objectif central de susciter un espace de discussion et d'échanges s'adressant aux collègues et aux étudiants de la discipline, mais pas seulement. En effet, les enjeux des transformations liées au numérique ne se réduisent pas à une approche

disciplinaire, mais suscitent un regard réflexif que les Sciences de l'Information et de la Communication gagneront à instruire et à construire. La quarantaine de collègues présent.e.s a partagé le constat que la transformation numérique de nos sociétés engendre un nombre grandissant de projets théoriques, épistémologiques, voire disciplinaires, tels que ceux portés par les Humanités Numériques et, plus récemment, les *Digital Studies*. Selon eux, cette effervescence de la réflexion sur le « numérique » ne va pas sans poser certains problèmes et rend pertinente une prise de position collective des enseignant.e.s-chercheur.e.s en information-communication, laquelle s'est exprimée dans une première publication coordonnée par Françoise Paquienséguy, au titre de sa vice-présidence recherche de la SFSIC.

Des échanges de points de vue au manifeste disciplinaire

Le groupe de chercheur.e.s[1] à l'origine de ce travail collectif a porté les bases de la réflexion issue de journées d'études en rédigeant un « *Manifeste pour un positionnement des Sciences de l'information-communication vis-à-vis des Digital Studies et autres mutations du numérique* ». Rapidement publié dans le dixième numéro de la *Revue Française des Sciences de l'Information et de la Communication* en décembre 2017 (https://rfsic.revues.org/2630), il réunit une vingtaine de prises de position d'enseignants-chercheurs de la discipline autour de trois grands thèmes : le numérique comme changement de paradigme ; la contribution des Sic à la genèse des Humanités Numériques (HN) ; enfin, la multiplicité des débats autour du transhumanisme. Sans être totalement satisfaisante, manquant quelque peu d'unité et d'homogénéité, cette œuvre collective et ouverte a cependant permis, au travers des textes courts qui la constituent de formaliser de multiples interrogations, mais aussi de certitudes, sur les liens existants entre les Sic, les Humanités Numériques et les Études Digitales. Ses contributeurs ont par ailleurs souligné l'importance d'une meilleure valorisation des atouts et des positions de la discipline des Sic, dans un contexte académique où bien des voix se font entendre pour revendiquer la paternité ou la propriété des recherches sur le « numérique »…

1. Groupe de réflexion constitué par Julia Bonaccorsi (Université Lyon2/ELICO), Nicolas Pélissier (Université Nice Sophia Antipolis/I3M), Louise Merzeau (Université Paris X/DICEN-IDF), Serge Bouchardon (Université de technologie de Compiègne/COSTECH), Valérie Carayol (Université Bordeaux Montaigne/MICA), Manuel Zacklad (CNAM/DICEN-IdF), sous la coordination de Françoise Paquienséguy (SciencesPo Lyon/ELICO).

Du manifeste au symposium : un double effort de formalisation

Suite à la publication du manifeste, une coopération inédite et féconde[2] a associé la SFSIC (présidée par Daniel Raichvarg puis Philippe Bonfils) et la Conférence Permanente des Directeurs.trices de Laboratoires en Sic (CPdirsic, présidée par Jacques Walter) dans l'objectif de mieux formaliser et donner un débouché éditorial plus large à la réflexion collective de ses auteurs, ainsi que de tous les chercheur.e.s en Sic se sentant concerné.e.s par le projet. C'est ainsi que tous les laboratoires membres de la CPdirsic ont pu, quand ils l'ont souhaité, désigner un de leurs membres pour participer aux deux symposiums qui ont permis la transition d'un Manifeste porteur de positions personnelles à l'ouvrage coécrit par un collectif d'une trentaine d'auteurs que ce texte introduit. Le travail de préparation et de coordination a été assuré en binôme par Nicolas Pélissier (pour la CPdirsic) et Françoise Paquienséguy (au titre de la SFSIC). Les participant.e.s se sont engagé.e.s à produire un travail collectif visant un consensus, afin de construire ensemble une prise de position disciplinaire quant aux défis posés par l'écosystème numérique. Les deux symposiums se sont tenus les 16 mars et 12 octobre 2018, dans les locaux de l'Institut du Management de l'Information[3] (Université Technologique de Compiègne) et ont de nouveau réuni plus d'une quarantaine d'enseignant.e.s-chercheur.e.s.

Les participant.e.s au symposium sont partis de plusieurs constats. En premier lieu, celui de l'intégration croissante des Humanités Numériques dans les parcours de formation, phénomène qui a également été mis en évidence par les « *Rencontres Humanités Numériques* » organisées en mars 2017 par la DGSIP du Ministère de l'Enseignement Supérieur et de la Recherche. En second lieu, les chercheur.e.s présent.e.s ont constaté l'importance accordée au « numérique » par les revues disciplinaires : pérennité des revues les plus anciennes telles que *Réseaux* ou *Terminal*, renouveau de *TIC et Sociétés*, succès non démenti des *Cahiers du Numérique* ou *Interfaces Numériques*, essor de nouveaux supports telles que *Études Digitales...*

2. Comme l'ont noté Daniel Raichvarg et Jacques Walter lors de l'ouverture du premier symposium, cette coopération a été la première de ce type, avec celle qui a conduit le Conseil National des Universités, la SFSIC et la CPdirsic à réfléchir ensemble à une reconnaissance conjointe des revues disciplinaires.

3. Nous remercions tout particulièrement Serge Bouchardon, Jérôme Valluy et les collègues du laboratoire COSTECH d'avoir accueilli à deux reprises notre symposium et l'ensemble des réunions du comté éditorial du présent ouvrage.

sans oublier la multiplicité des dossiers consacrés aux Humanités Numériques par des revues généralistes telles que la *Revue Française des Sciences de l'Information et de la Communication* (n° 8, 2016 ; n° 12, 2018) ou encore *Questions de Communication* (n° 31, 2017). Dans le même registre, mentionnons le nombre croissant de manifestations scientifiques interrogeant les relations entre Sic et Humanités Numériques, à l'image du 23e colloque franco-roumain en Sic (Cluj, octobre 2018), qui a débouché sur l'ouvrage collectif *Information, Communication et Humanités Numériques* publié chez Accent (Ioan Roxin, Nicolas Pélissier *et alii*, 2019) ou des colloques thématiques récurrents tels que EUTIC, TICEMED ou H2PTM.

Enfin, ces évènements et publications scientifiques ont soulevé un certain nombre d'interrogations que les participant.e.s au symposium ont souhaité structurer en quatre thématiques interreliées, qui se retrouvent dans les quatre parties du présent ouvrage. Chaque groupe de travail a cherché ainsi à répondre, dans le cadre d'ateliers dédiés, à un ensemble de questionnements débouchant sur une question centrale.

Une structuration de la réflexion autour de quatre questions fondamentales

L'écosystème numérique : un renouvellement théorique et épistémologique ?

Par rapport à d'autres disciplines plus anciennes, qui ont tendance à envisager le numérique comme une source de renouvellement, voire un impératif de survie, les Sic bénéficient d'une expérience non négligeable. Par exemple, les travaux des années 1980-1990 sur le triangle sciences-techniques-société, l'informationnalisation et les TIC ont permis aux Sic de penser le « numérique » avant son avènement contemporain, mais sans parvenir vraiment à remettre en perspective durablement ce terme discutable et discuté.

Pourtant, l'accent mis par les Sic sur les questions d'usages des TIC et sur la réception des médias et, plus récemment, les travaux de collègues en 71e section sur des domaines tels que les communs de la connaissance, les traces numériques, la protection des données et leur accessibilité, le capitalisme informationnel, la créativité digitale, les plateformes participatives, les corpus et le sens, ou encore les réseaux socionumériques les prédisposent légitimement à jouer un rôle central dans le dialogue avec les nouvelles interdisciplines ou postures que sont les Humanités Numériques et les *Digital Studies*.

La transformation numérique : des enjeux politiques et sociétaux à repenser ?

La position de la recherche en Sic face aux discours politiques renforçant les potentialités de contrôle, d'aliénation et d'assujettissement engendrées par la digitalisation des modes de communication, de pouvoir et d'organisation reste confuse. Pourtant, les chercheur.e.s en Sic gagneraient à mettre l'accent sur les nouvelles formes de « risques techno-sociaux » et à mobiliser à nouveaux frais des concepts tels que celui de « dispositif sociotechnique d'information et de communication » ou de « capitalisme informationnel » pour en penser les conséquences. En contrepoint, ils pourraient aussi proposer des recherches appliquées destinées à renforcer la dimension émancipatrice des technologies digitales (contre-récits au *storytelling* dominant, nouveaux outils participatifs et citoyens, nouveaux modes d'accès partagés au savoir et d'intelligence des territoires, etc.). Dans la logique du congrès de la SFSIC en juin 2018, les recherches portant sur la création et la créativité – en partenariat avec des spécialistes d'art, de design ou de management – pourraient d'ailleurs attester d'un engagement plus marqué dans cette voie de changement social et environnemental.

Les méthodes de recherche : nouvelles pratiques, nouveaux outils ?

Les outils de travail des chercheurs en Sic ont changé et se sont très vite démultipliés. Sans préjuger de ce constat, certaines méthodes semblent s'autojustifier par leur seule référence à leur caractère « numérique ». Il convient aujourd'hui de faire le point sur la valeur ajoutée et les limites de ces méthodes (questionnaire ou observation en ligne, entretien par Skype, etc.) par rapport aux méthodes traditionnelles antérieures. Plusieurs questions se posent aussi sur les corpus automatiquement constitués, sur la collecte de données par des robots ou sur les singularités des méthodes digitales dites « natives ». En banalisant ces méthodes, les Sic courent le risque d'une dérive « cartographique » et quantitativiste, mais aussi d'un usage exagéré de méthodes importées des sciences de l'ingénierie ou de gestion. D'où une interrogation sur la mise en place de comités d'éthique agissant a priori ou bien sur l'élaboration de chartes à partir notamment des exemples nord-américains ou scandinaves.

Les enjeux de formation : quel socle de compétences pour la littératie numérique ?

Les unités d'enseignement et les diplômes placés sous la bannière des « Humanités Numériques » ou « *Digital Studies* » ont récemment enrichi l'offre de formation en Sic. Ces formations font souvent état d'approches pluridisciplinaires et rapprochent les Sic de l'informatique et des méthodes « natives ». Le dilemme se fait jour entre deux postures : développer des formations spécifiques interdisciplinaires labellisées « Humanités Numériques », ou alors renforcer les enseignements sur l'écosystème numérique au sein des formations disciplinaires. Quel que soit le choix, la difficulté persiste d'actualiser les compétences des enseignants et la transmission des savoirs dans un domaine où le renouvellement des techniques de collecte et de traitement des données est de plus en plus rapide, et où la place accordée aux savoir-faire techniques et à la maîtrise logicielle est dominante… ou alors absente. Par ailleurs, les techniques de pédagogie dites « innovantes » et les « tiers-lieux » (*co-working*, fablabs, ateliers créatifs, etc.) sont promus par les acteurs du territoire et par un discours d'escorte généralisé… à tel point que les structures d'enseignement s'en dotent, progressivement ou à marche forcée, ce qui ne va pas sans poser certains problèmes de fond qui interpellent la recherche en Sic.

Au cours des deux sessions du symposium et autour d'elles, le travail a été intense et le rythme soutenu. Au total, près d'une trentaine d'auteurs ont rédigé quatorze chapitres, à une ou plusieurs voix, tout en formulant des positions et propositions émanant de leur communauté disciplinaire. Comme le lecteur le percevra sans doute, l'exercice n'a pas été toujours simple, et quelques glissements difficilement évitables se sont parfois produits entre ces chapitres. Mais le résultat est là, structuré autour de quatre parties, elles-mêmes subdivisées en trois chapitres chacune, correspondant aux subdivisions formulées par les participant.e.s lors du symposium.

La première partie de l'ouvrage traitera ainsi des opportunités de renouvellement théorique et épistémologique offertes aux Sic par les Humanités Numériques ; la deuxième se penchera davantage sur les enjeux politiques et sociétaux à l'œuvre en amont et en aval des programmes pédagogiques et scientifiques liés aux Humanités Numériques ; la troisième partie, quant à elle, abordera les questions de méthodologie, en interrogeant notamment les atouts et limites des méthodes dites « digitales » ; enfin, la quatrième et dernière

partie approfondira le lien entre formation et recherche sur les environnements digitaux, en réfléchissant plus particulièrement au rôle des Sic dans les modalités d'apprentissage, au-delà des outils, d'une littératie voire d'une citoyenneté numérique.

Ce résultat est le fruit du travail vigilant du *comité éditorial* de l'ouvrage[4], qui a pris en charge sa réalisation, depuis la conception et la mise en cohérence des textes, jusqu'à leur finalisation et l'édition-publication du présent ouvrage sous l'égide de la CPdirsic et de la SFSIC. Que les membres de ce comité en soient chaleureusement remercié.e.s !

Françoise Paquienséguy
et Nicolas Pélissier

4. Ce comité est composé de Gloria Awad (Université d'Artois) ; Christine Barats (Université Paris Est Créteil, CEDITEC) ; Ghislaine Chabert (Université de Savoie, LLSTI) ; Franck Debos (Université Côte d'Azur, Sic.Lab) ; Brigitte Juanals (Aix Marseille Université, CNE) ; Françoise Paquienséguy (SciencesPo Lyon, ELICO) ; Nicolas Pélissier (Université Côte d'Azur, Sic.Lab) ; Éric Triquet (Avignon Université, CNE).

PARTIE I

Vers un renouvellement théorique et épistémique ?

INTRODUCTION PARTIE I

Vers un renouvellement théorique et épistémique ?

L'objectif de cette partie est double. Il cherche à retrouver la filiation des Humanités Numériques en soulignant les points forts de sa généalogie. Par transfert des apports de Simondon sur la construction et l'essence de la lignée technique, il s'agira ici de retrouver l'essence des Humanités Numériques, en termes de disciplines, de concepts et de faits marquants qui en ont infléchi le cours. Mais dans un même élan, il cherche à mettre en lumière les apports des Sciences de l'Information et de la Communication à la construction de cette lignée, dont elle a compris les forces et les différences qui forgeaient sa singularité et sa complexité ; autrement dit sa richesse. Voilà pourquoi les six auteurs de cette partie, tous aux prises avec le numérique, ses humanités et ses technologies depuis l'origine, ont souhaité construire le raisonnement en trois temps, chacun confié à un chapitre.

Le premier, intitulé *Des enjeux épistémiques aux questions sociopolitiques* est écrit par Roger Bautier. Ce chapitre s'ouvre sur le constat et la certitude que les précurseurs des Humanités Numériques se sont saisis de l'écrit et du « perfectionnement des potentialités de l'écrit et des dispositifs graphiques » pour faire des propositions alliant l'analyse structurale ou le travail sur corpus et la mécanographie. Au fil du temps et des avancées, l'analyse des données conduit à une approche computationnelle qui dépasse les opérations de classement ou de tri au sein des corpus, plus ou moins grands. Ainsi l'informatisation des humanités alimente plusieurs questions dont celle du sens. Le sens donné à l'analyse quand les processus et procédures automatiques pèsent sur l'interprétation, le sens donné à et par ces outils de

traitement des objets sémiotiques si leur utilisation n'est pas adaptée au corpus analysé ; autrement dit deux conceptions des Humanités Numériques s'opposent. D'une part celle qui valorise l'informatique comme outil de traitement et d'analyse sur la base d'une manipulation formelle d'éléments et de l'autre, celle qui valorise les compétences interprétatives et analytiques du chercheur, appliquées à la fois aux données traitées et aux processus de traitement qui leur sont appliqués.

Le deuxième, intitulé *Les Humanités Numériques quelles définitions ?* est écrit par Dominique Carré, Fanny Georges et Jérôme Valluy. Ce chapitre cherche à proposer une définition des Humanités Numériques à partir des deux termes du syntagme, c'est-à-dire en cernant à la fois les humanités et le numérique mais surtout les tensions et les relations qui surgissent entre eux. S'appuyant sur les disciplines qui ont pensé l'humanisme, il saisit les humanités principalement portées par le système éducatif, pour le confronter au numérique de ce siècle dans lequel « la compétence informatique ne forme plus l'usage ». Autrement dit, devenues numériques ou digitales, les humanités contemporaines s'inscrivent dans une généalogie qui les positionne en « technologies du texte et des savoirs ». Plusieurs interprétations en découlent et le chapitre en discute principalement trois.

Tout d'abord celle qui fait des Humanités Numériques un label institutionnel de revalorisation, mais aussi un impératif de collaboration qui engendre des méthodes de recherche assises sur le numérique ; la conséquence majeure en est « la transformation des pratiques de recherche et de l'éthos du chercheur ». Ensuite, celle qui porte à une définition large des Humanités Numériques voyant s'entrechoquer les versions issues de la littérature anglophone et francophone en discussion avec celle du célèbre Manifeste des *Digital Humanities* de 2010 qui s'en veut le fondateur ; la place des technologies numériques devient alors centrale, voire première selon les positions disciplinaires des SHS. Au-delà des divergences, une définition élargie, parfois transformée en outil d'action ou de légitimation laisse la porte ouverte à l'échange et la construction pluridisciplinaire. Les Sic paraissent tout à fait outillées pour conduire, ou accompagner pour le moins, ces avancées. Six axes sont proposés pour guider les débats et les actes.

Le troisième, intitulé *Les Humanités Numériques depuis les Sic* est écrit par Chérifa Boukacem-Zeghmouri et Françoise Paquienséguy. Ce chapitre met en valeur les jalons posés par la discipline au fil des rapports de plus en plus étroits entre social, numérique et

technologies, data incluses, pour voir l'évolution des problématiques de la discipline au contact de celles plus largement portées par les Humanités Numériques. Pour ce faire, il s'appuie fortement sur trois éléments qui la caractérisent et l'outillent pour travailler les Humanités Numériques. Tout d'abord, les Sciences de l'information et de la communication sont une interdiscipline. Initialement, si une partie des concepts et des outils de la jeune discipline provient donc des disciplines antérieurement établies, les Sic ont développé une approche spécifique de l'information et de la communication pour en questionner et saisir les processus, le sens, les acteurs et leurs productions dans un mouvement d'ensemble. Ensuite, elles ont vu naître et se déployer le numérique, ses acteurs et ses technologies dont elles se sont immédiatement saisies, forgeant de nouveaux concepts pour les étudier puis pour les penser. Ainsi, les Sic témoignent d'une posture critique caractéristique de la distanciation qu'elle opère au regard de la technique et de la primauté du social dans ses approches. Posture épistémologique nécessaire pour comprendre et analyser les arrangements, les agencements, les structurations des réseaux, des acteurs, des dispositifs qui engendrent le passage : du réel à la connaissance, du fait à la représentation, de l'imaginaire à la technique, de l'objet à l'usage et de la data à l'information. Enfin, fortes d'une expérience à étudier les rapports entre le social et la technique depuis presque cinquante ans autour de plusieurs notions emblématiques qui sont présentées dans le chapitre et d'une légitimité scientifique que les deux termes information et communication assoient, la discipline travaille toujours à apporter des réponses à ces transformations liées au numérique, comme elle ne cesse de le faire depuis les années 1980.

Françoise Paquienséguy

CHAPITRE I

Des enjeux épistémiques aux questions sociopolitiques

Roger Bautier
Professeur émérite
Université Sorbonne Paris Nord, Labsic

Résumé : L'usage des méthodes computationnelles dans les disciplines qui appartiennent au domaine des lettres et des sciences humaines et sociales a déjà suscité de larges débats, mais, présents depuis le milieu du XXe siècle, ces débats ont été relancés par l'informatisation généralisée et par le développement du web depuis les années 90. Qu'ils se rattachent aux Humanités Numériques, qu'ils préfèrent d'autres dénominations ou qu'ils choisissent seulement de tirer les conséquences de l'essor du numérique, il s'agit, pour les chercheurs, de se positionner par rapport à des questions difficiles : comment envisager la place du sens dans les traitements informatiques et dans les résultats des analyses effectuées ? comment se comporter face aux déstabilisations affectant les disciplines ? comment traiter les implications politiques des nouvelles tendances de la recherche ?

Mots-clés : méthodes computationnelles, Humanités Numériques, place du sens, disciplines, implications politiques.

Abstract: The use of computational methods in disciplines belonging to the humanities and social sciences has already given rise to broad debates, but these debates, which have been present since the mid-twentieth century, have been revived by the widespread computerization and the development of the web since the 1990s. Whether they are related to the digital humanities, whether they prefer other denominations, or whether they simply choose to draw the consequences of the digital boom, researchers must position themselves in relation to difficult questions: how to envision the place of meaning in computer processing and in the results of the analyses carried out? how to deal with the destabilization affecting the disciplines? how to deal with the political implications of new research trends?

Keywords: computational methods, digital humanities, place of meaning, disciplines, political implications.

Deux références historiques apparaissent très fréquemment dans la présentation de l'histoire des Humanités Numériques : Roberto Busa (avec les travaux qu'il effectue à partir de 1946) et la *Text Encoding Initiative* (signée en 1987). On évoque alors l'époque du *literary and linguistic computing* et on signale qu'elle a été suivie par celle du *humanities computing*, elle-même suivie par celle des *digital humanities*. Mais il est sans doute nécessaire de replacer ces jalons dans l'ensemble de l'histoire du traitement de l'information depuis l'invention de l'écriture jusqu'à l'essor du web, en passant par l'invention de l'imprimerie et ses conséquences en matière de diffusion de l'information (Le Deuff, 2018). Comme l'a montré notamment l'anthropologie de l'écriture, les changements dans les modes de communication ont une incidence sur le développement des processus cognitifs et sur l'accroissement du savoir. L'écriture, de ce point de vue, représente la possibilité pour l'esprit critique de s'exercer plus facilement, dans la mesure où elle déploie le discours devant les yeux et où elle autorise une meilleure accumulation des connaissances en permettant à la communication d'échapper aux limites du simple contact personnel. Plus globalement, avec l'écrit proprement dit et avec les dispositifs graphiques (listes, diagrammes, tableaux), émerge la possibilité de prendre du recul par rapport au discours et de l'examiner d'une manière plus abstraite et plus rationnelle, une possibilité qui semble capable de rendre conscient des différences en général et des contradictions en particulier, ainsi que de pousser à la réflexion sur les règles de raisonnement.

En fait, c'est dans le perfectionnement des potentialités de l'écrit et des dispositifs graphiques que les précurseurs directs des Humanités Numériques se sont engagés. En particulier, on peut trouver ces précurseurs dans toutes les disciplines des sciences humaines et sociales qui ont été le lieu de l'expérimentation de l'analyse structurale, une expérimentation qui a porté, dès les années 50 et 60 du XX^{e} siècle, sur des corpus plus ou moins grands. C'est ainsi qu'un lien s'est établi

rapidement entre l'analyse structurale en anthropologie et le recours à la mécanographie (qui préfigure le traitement de l'information sur ordinateur). Le volume des textes mythiques qui sont étudiés entraîne des contraintes : si le système de référence comprend plusieurs dimensions, il sera nécessaire d'utiliser des cartes perforées, l'usage d'un simple stock de fiches n'étant évidemment plus suffisant. Pour l'analyse automatique du discours qui apparaît aussi au cours de cette période, la problématique structuraliste sert à combattre les études les plus courantes, notamment l'herméneutique littéraire, avec une conception du sens liée à la conscience du lecteur, mais également l'analyse de contenu pratiquée par les sciences humaines. Au même moment, on met l'accent sur la nécessité de constituer des répertoires exhaustifs de structures syntaxiques afin de parvenir à une meilleure analyse linguistique des textes. L'essor du traitement automatique de l'information en archéologie est un autre bon exemple : lorsque les données sont trop nombreuses ou trop complexes pour être traitées d'une autre manière, les ordinateurs peuvent sortir de la fonction de calcul et de simple tri pour fournir un outil plus précieux. L'objectif est alors qu'ils servent, dans un premier temps, à classer des données mais, surtout, dans un second temps, à en tirer, suivant une démarche relevant d'un calcul au sens logique (et non arithmétique), une vision globale permettant de maîtriser tout un ensemble de relations entre elles.

Les approches computationnelles dans les sciences humaines et sociales sont donc déjà anciennes et, dans beaucoup de secteurs qui se sont développés depuis ces deux décennies, les chercheurs ne considèrent pas obligatoirement leur travail comme relevant des Humanités Numériques. Dans certains cas, il est même remarquable que les approches adoptées soient plus proches de celles des sciences dures que des projets rattachés aux Humanités Numériques. Quant aux humanités au sens classique, il est possible d'y repérer une opposition entre deux attitudes. Il semble, en effet, que celles-ci aient été, depuis leur origine, marquées certes par une attitude bien connue consistant à étudier ce qui est singulier, voire exceptionnel, mais aussi par une autre, qui met l'accent, au contraire, sur les régularités. Si une telle opposition existe et paraît plus pertinente, les Humanités Numériques sont assimilables à un prolongement instrumenté de cette dernière attitude. Ne s'agit-il pas simplement d'admettre que comprendre la société impose d'utiliser toutes les technologies de traitement de l'information, y compris la plus récente ? Après l'écriture et après l'imprimerie, l'informatique et les technologies

du web constituent des outils sans doute incontournables. Ce n'est certes pas une raison pour négliger les débats que les Humanités Numériques peuvent susciter, qu'il s'agisse de la place du sens dans les analyses, des relations entre les types de recherche ou du danger des dérives éventuelles.

Le sens en question

La rupture (que beaucoup de chercheurs assimilent à une rupture épistémologique) introduite par l'informatisation des humanités ne s'identifie pas systématiquement à une adoption du modèle des sciences de la nature, mais elle correspond bien à une transformation des conditions de réalisation du processus d'interprétation qui continuent de marquer les humanités, par le recours, notamment, à des procédures de traitement automatique de grands corpus numérisés. Un des risques est ainsi que les méthodes informatisées d'analyse d'objets sémiotiques réduisent considérablement le champ de l'examen de ces objets. En même temps, il est nécessaire de se demander dans quelle mesure les dangers de l'interprétation (qui serait trop subjective) sont effectivement écartés par les méthodes informatisées. Si l'on veut conserver l'orientation consistant à s'intéresser à la construction du sens dans les corpus, une des possibilités existantes est de donner toute son importance, justement, au corpus (Longhi, 2017). Cela revient à privilégier son analyse strictement interne, en lui donnant le rôle de pourvoyeur des éléments permettant d'interpréter l'ensemble de ses composants. S'ajoute à cela la question de savoir si les résultats de l'utilisation d'une procédure informatique sont adaptés aux propriétés des objets étudiés (Meunier, 2017). Par exemple, un résumé automatique sera peut-être d'un faible intérêt lorsqu'il portera sur une œuvre littéraire. De façon complémentaire, se pose la question de savoir si les interprétations qu'on est amené à donner sont bien fondées : un nuage de mots, par exemple, qui n'est que la visualisation d'un certain traitement informatique du texte, peut donner lieu à des références à des champs conceptuels multiples, références toujours discutables.

Les oppositions classiques entre qualitatif et quantitatif, interprétatif et empirique, etc., sont présentes dans les discussions sur les Humanités Numériques. Cependant, le trait le plus décisif qu'il faut rappeler au sujet du numérique, c'est que celui-ci, en se rattachant au calcul, n'a besoin que de la manipulation formelle d'éléments sans signification, ce qui situe son domaine aux antipodes de l'herméneutique (Bachimont, 2010). Les Humanités Numériques

peuvent être vues comme le résultat d'un recours à l'informatique en tant qu'outil permettant de mieux analyser les objets étudiés. Mais elles peuvent également être envisagées comme impliquant des données que les procédures informatiques vont traiter pour donner des résultats que le chercheur devra interpréter. Une opposition tout aussi forte apparaît au niveau de la conception de la place du sens. D'un côté, les objets étudiés sont fondamentalement porteurs de sens, alors que, de l'autre, il s'agit de travailler à fournir d'eux une représentation formelle qui débouchera sur la possibilité d'une validation empirique. Des ruptures radicales apparaissent ainsi, entre les données et leur origine, entre la simplicité de la présentation des résultats et la complexité des traitements, entre la diversité des données et leur traitement unifié. L'incertitude qui pèse sur la fabrication de ces données fait que le calcul qui porte sur elles pose le problème de la construction du sens (ainsi, le statut de trace accordé à un objet n'est pas issu de l'objet lui-même, mais est établi par l'activité du chercheur). Le langage n'est, dès lors, plus central et, comme les sciences expérimentales l'ont fait pour la connaissance de la nature, les Humanités Numériques cherchent à produire la connaissance par le calcul, avec des avantages, certes, mais aussi avec des inconvénients.

L'accumulation de données paraît primordiale, mais que faire de celles-ci lorsqu'elles ont été accumulées ? Une idée profondément nouvelle a émergé : l'idée que tout corpus peut être parcouru, que les enquêtes peuvent parvenir à l'exhaustivité et que le document est en fait le document numérique, le numérique apparaissant comme une nouvelle forme du corpus. Il faut souligner que l'instrumentation numérique de la recherche implique des procédures techniques qui renvoient à la problématique des rapports entre écriture et modes de pensée. Il en résulte que les Humanités Numériques mettent en œuvre des dispositifs qui peuvent être analysés comme concourant à la visualisation des connaissances, ce qui renvoie à une sémiotique de l'écriture, autant qu'à des considérations proprement techniques ou relevant de la sociologie des sciences. C'est pourquoi, pour certains chercheurs, la lecture automatique des corpus peut sans doute s'effectuer dans une perspective plus herméneutique. Dans cette direction, il s'agirait de perfectionner la visualisation des données, afin de favoriser l'interprétation des grandes masses de données. Mais doit-on toujours penser la culture comme des données pouvant être explorées et visualisées ? La visualisation des réseaux, par exemple, s'est considérablement amplifiée au cours des dernières années au sein des Humanités Numériques et des sciences sociales dans leur ensemble.

Elle est censée, en effet, permettre une meilleure saisie intuitive des résultats de l'analyse (ce qui ne signifie pas qu'elle puisse servir de preuve). Il faut cependant souligner que, en général, les algorithmes à la base des visualisations de réseaux produisent une disposition des nœuds qui correspond non pas à une projection des réseaux dans un espace de visualisation, mais, en fait, à une spatialisation, c'est-à-dire à la création d'un espace qui n'existerait pas sans ces algorithmes (Venturini, Jacomy, Jensen, 2019).

C'est notamment dans la mesure où elle donne le moyen de privilégier les méthodes qui suspendent la compréhension immédiate d'un texte que l'informatique apparaît aisément comme un garant méthodologique. Les nouvelles relations entre la linguistique et l'informatique montrent que l'annotation syntaxique, y compris automatique, laisse la place à des traitements de chaînes de caractères, d'où des situations fort différentes : le traitement automatique des langues vise la performance technique quand l'analyse linguistique cherche à améliorer la connaissance du fonctionnement du langage. La traduction automatique en constitue une bonne illustration, puisqu'elle a bénéficié d'une modification radicale de méthode qui dépend directement de cette évolution (Poibeau, 2014). Au lieu de s'intéresser aux problèmes posés par les ambiguïtés propres à toutes les langues et par l'importance des connaissances sur le monde qui sont nécessaires à la compréhension, elle a fait des progrès décisifs en se fondant sur des alignements établis à partir de corpus bilingues. D'où la possibilité de trouver des équivalents entre deux langues par une analyse automatique travaillant sur la statistique des séquences de mots. Il n'y a pas besoin de modéliser la sémantique d'un texte pour le traduire : on a plutôt un assemblage automatique de morceaux de traductions déjà réalisées. La mise en suspens de la compréhension immédiate est tout aussi visible dans l'analyse des données culturelles à une grande échelle (Manovich, 2017). Cette dernière a porté d'abord sur les archives numérisées créées par les bibliothèques et les universités. Dans un deuxième temps, ce sont les contenus culturels présents sur le web, ainsi que les conversations et les commentaires qui en sont le prolongement, qui doivent être étudiés, si l'on veut comprendre les évolutions culturelles d'aujourd'hui (et, éventuellement, en tirer des enseignements pour l'étude des productions du passé). Il s'agit alors d'élaborer des *cultural analytics* (on parle aussi d'humanités computationnelles). Les *cultural analytics* sont vues comme intégrant des préoccupations communes avec, à la fois, les Humanités Numériques et les travaux d'informatique sociale

(*social computing*), tout en s'en différenciant : elles s'intéresseraient plus à la culture numérique contemporaine que les premières et plus aux aspects culturels des réseaux sociaux que les seconds. Ce type de préoccupation entraîne la délinéarisation du texte car les outils utilisés parcourent tout le corpus et produisent des résultats le plus souvent sous forme de listes, ce qui permet, par exemple, de repérer des éléments récurrents qui, autrement, seraient imperceptibles sur de grands corpus (Moretti, 2016). En effet, la base de données rend compte du monde sous la forme d'une liste d'éléments. La connaissance ne va donc pas reposer sur la lecture de textes, mais sur le traitement de corpus (constitués de textes). La lecture porte sur des objets qui ont été écrits pour entrer dans un processus de production et de transmission de sens, alors que, lorsqu'il s'agit de corpus, on a affaire à des objets créés par des chercheurs, objets qui n'entrent pas dans ce processus. Cette critique computationnelle est ainsi amenée à contourner le sens, au moins à un premier niveau, mais les algorithmes fournissent des résultats abstraits qui permettent de fonder des constructions intellectuelles dont on pourra se servir pour enrichir l'analyse des textes. C'est pourquoi il n'est pas incongru de parler d'une véritable herméneutique numérique (Mayaffre, 2007), en admettant la complémentarité entre la lecture linéaire (qualitative) et la lecture tabulaire et réticulaire (quantitative), qui correspondent à des interrogations différentes.

Des disciplines déstabilisées

Le développement des Humanités Numériques pose la question des relations avec l'informatique, ainsi qu'avec les sciences humaines et sociales, mais aussi avec les sciences de la nature. On peut, bien sûr, juger que l'opposition entre sciences dures et sciences molles doit être dépassée, grâce à la prise en compte d'une globalité intellectuelle, qui serait marquée par une rationalité dépassant le cadre du calcul. Pour certains, au contraire, l'option préférée consiste à considérer qu'il faudrait parler de sciences numériques plutôt que d'Humanités Numériques. C'est ce qui explique probablement que l'on parle quelquefois d'Humanités Numériques et scientifiques, ce qui peut impliquer, en fait, une intégration des humanités aux sciences exactes par l'intermédiaire de l'informatique. La plupart du temps, l'attention est centrée sur la transformation des humanités traditionnelles, qu'il s'agisse des humanités littéraires ou des humanités au sens anglo-saxon, c'est-à-dire renvoyant aussi bien aux sciences humaines et sociales qu'aux humanités littéraires, et l'objectif de nombreux

commentateurs est d'évaluer les conséquences de la transformation entraînée par l'usage de l'informatique (Bigot, 2018). Si la numérisation de la littérature primaire et secondaire de la recherche paraît un aspect fondamental, il reste que la réflexion sur les Humanités Numériques se penche préférentiellement sur l'apparition de nouveaux objets et de nouveaux cadres de recherche sur ces objets, qui sont liés à l'introduction de méthodes nouvelles. En même temps, elle peut mettre en lumière l'interdisciplinarité qui semble être une caractéristique des Humanités Numériques ou, plus exactement, un de leurs effets. Dans ce cas, on envisage le numérique comme le lieu d'un dialogue nécessaire entre disciplines. Par ailleurs, un certain flottement, au niveau épistémologique aussi bien qu'au niveau institutionnel est repérable. Par exemple, l'utilisation du terme « humanités » peut signifier la référence à des textes et à des documents, ce qui peut se concrétiser par une centration sur la matérialité linguistique ou sémiotique des objets considérés, mais pas obligatoirement. En fait, la perspective de bouleversements dans l'organisation de la recherche universitaire est fréquemment associée au déploiement des Humanités Numériques, dans un contexte qui donne la prépondérance aux sciences dures (Mounier, 2018). On doit remarquer, cependant, que, si ces dernières sont marquées par le passage d'une méthode expérimentale hypothético-déductive à un usage massif de données dont on cherche à tirer des tendances, il n'est pas sûr que les Humanités Numériques soient en train de se rapprocher du modèle classique des sciences dures : ne se fondent-elles pas plutôt dans une évolution générale caractérisant l'ensemble des sciences ? Inversement, au sein des sciences humaines et sociales, les usages du numérique sont manifestement très variés. Il est tentant de se demander si, dès lors, les sciences de l'information et de la communication ne pourraient pas devenir une sorte de discipline transversale. Celle-ci entrerait alors en concurrence avec la linguistique, qui apparaît de plus en plus comme tournée vers la production de données, aux dépens de l'exercice de sa fonction d'analyse, ce qui lui donne le rôle de fournisseuse de méthodologies au sein des Humanités Numériques.

On peut penser que, dans beaucoup de sciences humaines et sociales, la pratique scientifique passe par le recours au numérique. Pour celles qui se rattachent plus ou moins aux Humanités Numériques, trois attitudes existent : prolonger les comportements anciens (autrement dit, éviter le numérique), adopter l'ensemble des nouvelles procédures (avec leurs conséquences en termes d'exploitation systématique de données) ou chercher à établir la place des sciences sociales dans

ce nouvel environnement intellectuel (Boullier, 2015). Lorsqu'il est possible de trouver des objets d'étude permettant la collaboration interdisciplinaire (y compris lorsque la contribution d'internautes vise à contribuer à l'élaboration de connaissances nouvelles), les méthodes computationnelles utilisées dans divers secteurs de la recherche apparaissent comme des tentatives de rapprochement des humanités, conçues comme marquées par la subjectivité, et des sciences exactes, marquées par la prétention à l'objectivité (Cocco *et al.*, 2018). La cohésion peut se réaliser ou tenter de se réaliser par l'intermédiaire de valeurs partagées et mises en lumière dans un certain nombre de manifestes qui ont accompagné l'épanouissement de différents courants de recherche. On constate que, d'une certaine manière, les Humanités Numériques reprennent des thématiques qui relevaient des sciences de l'information, mais elles introduisent de nouvelles orientations grâce aux possibilités nouvelles de l'analyse quantitative des objets numériques, ce qui peut entraîner des prolongements au-delà des Humanités Numériques proprement dites. C'est pourquoi, dans le contexte du développement des bibliothèques numériques, par exemple, l'intelligence artificielle peut être envisagée comme un prolongement des outils déjà connus que procurent les Humanités Numériques, un apport prometteur particulièrement dans le domaine de la reconnaissance des formes ou dans celui de l'analyse des images.

Que l'on parle d'Humanités Numériques, d'humanités digitales, d'*humanities computing*, de linguistique computationnelle, de *digital studies*, de *digital methods*, de *cultural analytics*, voire de *computational communication science*, il est toujours tentant de souligner les transformations qui sont introduites dans des disciplines plus ou moins anciennes, pour lesquelles cet usage nouveau représente une aubaine aux yeux de certains chercheurs, alors que d'autres le jugent inintéressant ou négatif. Les critiques primordiales viennent de ceux qui soupçonnent les Humanités Numériques de se réclamer d'un mouvement de scientifisation des sciences humaines et sociales dont l'un des effets les plus nets serait d'établir une nouvelle grille d'évaluation des recherches attribuant la valeur la plus élevée à celles qui ont recours aux méthodes numériques. Il en découle d'autres critiques, qui, au-delà des questions épistémologiques posées par ces méthodes, regardent vers l'aspect politique des Humanités Numériques : celles-ci peuvent apparaître comme cherchant à imposer un modèle de savoir marqué par la référence à des normes très discutables de productivité des connaissances, plus ou moins liées à des intérêts industriels ou à des perspectives de contrôle social.

Cependant, les partisans des Humanités Numériques paraissent surtout mettre en avant les avantages procurés par les méthodes numériques lorsqu'elles sont appliquées à leur domaine de recherche, que celui-ci relève plutôt des humanités au sens classique du terme ou, plus largement, des sciences humaines et sociales (Valluy, version du 27/01/2020). Ceux qui défendent plutôt l'idée qu'il faut se consacrer aux études digitales (ou *digital studies*) auront tendance à prôner l'étude des conséquences globales de l'expansion du numérique sur la société. Il faut remarquer, par ailleurs, que les Humanités Numériques présentent deux orientations. D'une part, il peut s'agir, certes, de cerner le rôle des outils numériques dans les humanités ; d'autre part, il peut s'agir aussi de considérer le tournant numérique du point de vue des humanités. Des chercheurs de différentes disciplines peuvent se retrouver dans l'un ou dans l'autre de ces courants, quelquefois en fonction de leur proximité plus ou moins grande avec le domaine des humanités classiques, quelquefois pour des raisons stratégiques (en relation avec les financements de la recherche). D'autres encore ne s'y retrouvent pas, alors même qu'ils ont une pratique régulière, éventuellement intensive, des méthodes numériques au sein de telle ou telle des sciences humaines et sociales. Quant aux sciences de l'information et de la communication, elles paraissent partagées : on trouve, en leur sein, aussi bien des formations aux métiers des bibliothèques et de l'information scientifique et technique, renvoyant plutôt à la première orientation, que des recherches fortement pluridisciplinaires qui se rattachent plutôt à la seconde. Dans ces conditions, il est bien compréhensible que des stratégies académiques se dessinent, qui font apparaître des clivages ou des regroupements intellectuels nouveaux, couplés à des distinctions professionnelles inédites.

Un reflux du politique ?

Dans un premier temps, les humanités ont fait appel à l'informatique afin de traiter des documents initialement non numériques : la structure de ces documents sert alors de référence pour l'organisation des données dans lesquelles l'information est présentée numériquement. Le deuxième temps est celui de l'utilisation de bases de données. Celles-ci sont souvent rejetées, au motif que tout texte est marqué par une linéarité intrinsèque et par une certaine liberté d'expression, deux caractéristiques qui entreraient en contradiction avec celles des bases de données. Pourtant, il n'est pas impossible de considérer que la compréhension d'un lecteur, notamment celle d'un chercheur,

engendre la formation d'une représentation mentale qui transcende la séquentialité des différents éléments concrets du texte, d'où l'intérêt de mettre en lumière les différentes manières de représenter les composantes fondamentales du texte et d'établir leurs relations. Le web sémantique peut sembler représenter une avancée dans cette direction (Schloen, Schloen, 2014), puisqu'il traduit le passage d'un web de documents à un web de données, par l'intermédiaire d'une représentation standardisée des connaissances : les triplets sujet-prédicat-objet du langage RDF (Resource Description Framework), ainsi que les micro-formats et les micro-données. Ces recherches sur le web sémantique se donnent clairement pour objectif de parvenir à une utilisation perfectionnée des ressources du web par les humains et par les machines : on vise la mise au point d'un programme d'interopérabilité entre toutes sortes de données qui postule, d'une part, l'indépendance entre les connaissances représentées et les langues dans lesquelles elles sont produites et transmises, d'autre part, leur représentabilité systématique par un formalisme logique. Il s'agit donc de recourir à des ontologies, qui conceptualisent un domaine de connaissance par définition des concepts et de leurs relations, sous une forme qui peut donner lieu à des traitements automatiques (les ontologies sont encore relativement peu présentes au sein des humanités, alors que la massification des données y est évidente).

Le web sémantique peut-il fournir aux Humanités Numériques des moyens pour élaborer une représentation des données qui soit plus orientée vers l'expression du sens ? Cet objectif n'est atteignable qu'à la double condition que les informaticiens s'engagent dans la prise en compte des contextes dans le traitement des données au sein des Humanités Numériques et que les spécialistes des humanités et des sciences humaines et sociales, de leur côté, acceptent de préciser systématiquement la signification des éléments constitutifs du web sémantique qui peut leur être utile. Ce qui implique non seulement un travail sur les vocabulaires, mais aussi un travail sur l'explicitation des raisonnements qui se trouvent au fondement des interprétations. Avec une telle organisation des connaissances, ce qui pourrait être envisagé, c'est une formalisation capable de susciter la confrontation entre différentes recherches, soit sur le mode d'une harmonisation, soit sur celui d'une différenciation, éventuellement créatrice d'une transformation des positions dans le champ scientifique. Cependant, si l'on examine les politiques de numérisation, on observe que la recherche de l'ouverture et de l'interopérabilité des données est chose délicate (Juanals, Minel, 2017). La re-documentarisation à visée universaliste

qui est souvent prônée par les défenseurs du web sémantique est nécessairement limitée par des stratégies locales. Corrélativement, au cas où le passage à un web de données est jugé favorablement, il peut être intéressant de le compléter par le développement de formats d'annotation : si ceux-ci permettent d'annoter la ressource, ils autorisent aussi l'expression de réactions de l'usager et, par conséquent, le partage et la comparaison des interprétations. D'où un double mouvement envisageable : un éloignement par rapport au web sémantique et, au contraire, une avancée vers un web herméneutique, dont la composante calculatoire serait seulement au service de la confrontation des interprétations (Puig, 2017). Plus radicalement, il est possible de considérer que l'orientation souhaitable est celle qui consiste à s'intéresser à la sémantique du web plutôt qu'au web sémantique (Rastier, 2008). C'est ce que préconisent certains chercheurs en linguistique de corpus, qui pensent qu'il n'est pas sûr que l'on puisse faire la liste des objets qui composent le monde, que les connaissances sont sans doute plus inscrites dans des textes que dans les termes d'une ontologie et qu'il est douteux qu'elles soient entièrement représentables par un formalisme logique.

Par là même, l'examen des méthodes numériques débouche sur une réflexion concernant les aspects les plus discutables des effets de la mécanisation du sens. Cet examen incite d'abord à relativiser l'opposition entre analyse qualitative et analyse quantitative. Les méthodes numériques peuvent, en effet, favoriser leur complémentarité quand le statistiquement significatif se révèle non signifiant et le signifiant statistiquement peu significatif (Compagno, 2017). Un tel éclairage montre que les outils informatiques, qui n'ont pas accès au contenu, fournissent cependant une contribution précieuse à une analyse quali-quantitative des données, par leur capacité à transformer les formulations de ce contenu. On admet, dès lors, que cette transformation aboutira à des résultats qui, quelle que soit leur présentation, seront à la fois utiles et interprétables, au contraire des corpus d'origine, bien trop vastes pour être analysables directement par un humain. Il en résulte pourtant la nécessité d'une attention extrême aux modalités de la transformation de ces formulations. Cette attention est d'autant plus nécessaire que les méthodes numériques présentent des avantages incontestables, notamment en matière de quantité de données traitées et de visualisation des résultats. Ces avantages ne doivent pas faire oublier, d'une part, que les traitements effectués ne sont pas toujours clairement explicités et que leurs utilisateurs ne sont pas toujours capables de les comprendre, d'autre

part, que les visualisations plus ou moins spectaculaires risquent de se constituer en représentations d'autant plus difficiles à interpréter que les traitements eux-mêmes ne sont pas parfaitement maîtrisés. On remarquera que, si des critiques à l'égard des gouvernements et des entreprises qui cherchent à contrôler le traitement de l'information sont fréquemment formulées, cette vigilance n'a pas d'équivalent en ce qui concerne les rapports de pouvoir qui peuvent marquer les composantes mêmes de la production de sens.

Les conséquences d'un tel effet de « boîte noire » ne sont certainement pas négligeables et elles sont majorées par les caractéristiques des résultats que fournissent les méthodes numériques. Ceux-ci, en effet, conviennent aux critères d'évaluation issus d'un modèle universitaire néolibéral et, en se prêtant à une réutilisation par les industries culturelles, ils rendent plus facile l'établissement de liens avec le marché, ce qui incite à écarter des programmes de recherche des questionnements sociopolitiques pourtant bien ancrés dans les humanités et les sciences humaines et sociales, pour favoriser des démarches qui ne passent pas par ces questionnements. Plus largement, le risque se présente de voir ces méthodes favoriser une vision de la société très peu sociologique car les données traitées sont souvent loin de permettre la saisie de la réalité sociale (Ouakrat, Mésangeau, 2016). Donner toute son importance à la sémiotique est tout à fait compréhensible et souligner que l'étude des conditions de production de l'information implique, elle aussi, une analyse de données constitue un rappel utile, à condition, cependant, que les informations d'ordre social venant de l'environnement des ensembles de données soient effectivement utilisées, ce qui n'est pas toujours le cas (et, même, quelquefois, condamné). La difficulté à intégrer ce type d'informations semble, de surcroît, augmentée par le fait que les tentatives de naturalisation du sens ne sont maintenant pas rares et qu'elles orientent la recherche vers l'application au social d'approches qui sont celles de la biologie et de la physique. Ces tentatives vont dans le sens de l'établissement d'une science qui se veut totale, en rassemblant les différentes disciplines et en les transcendant grâce à l'application de quelques principes unificateurs. La « physique sociale », à laquelle certains courants des Humanités Numériques ne sont pas insensibles, repose ainsi sur l'idée que l'humain et la nature qui l'entoure ont en commun d'être soumis à des mécanismes computationnels et que les comportements des individus sont fortement prévisibles, ce qui permet d'envisager l'orientation de ces comportements afin de garantir leur conformité avec ce qui est

jugé souhaitable (Broca, 2016). Les travaux rattachés à la « physique sociale » n'ont pas tous cette caractéristique, mais ils habituent à considérer que l'accumulation des données numériques est toujours profitable et que celles-ci, compilables dans un tableau général des comportements individuels et sociaux, offriraient la potentialité de transformer au mieux la compréhension et la direction de nos vies, de nos organisations et de nos sociétés.

Des incontournables à surveiller

L'essor des Humanités Numériques et des différents courants de recherche qui s'y rattachent plus ou moins directement, d'une manière implicite ou explicite, se situe dans le prolongement de la longue histoire des moyens mis en œuvre pour analyser des objets sémiotiques. Bien évidemment, les possibilités offertes par l'informatique sont telles que des travaux jadis inimaginables sont maintenant réalisables et concourent effectivement à des progrès décisifs dans la connaissance de ces objets. Comme ces progrès dépendent de méthodes qui transforment profondément les processus d'interprétation, les chercheurs ne peuvent qu'être attentifs aux bouleversements inévitables, quoique potentiellement dangereux, de leurs pratiques et se positionner par rapport à des stratégies qui sont autant professionnelles que proprement intellectuelles. Ils sont amenés ainsi à devoir se pencher non seulement sur le renouvellement des problématiques du sens et sur celui des disciplines, mais également sur la vision de la société que – si l'on n'y prenait garde – ce renouvellement pourrait favoriser.

POSITIONNEMENT

– Les Humanités Numériques s'inscrivent dans la longue histoire des technologies de traitement de l'information. Celles-ci sont maintenant marquées par le développement de l'informatique, qui procure des outils incontournables.

– Il s'agit d'outils qui permettent, notamment, une mise en suspens du sens des objets analysés. Il en résulte des avantages et des inconvénients : l'objectif doit être de maximiser les premiers et de minimiser les seconds.

– Le ralliement ou non aux Humanités Numériques dépend largement de stratégies au sein des organismes de recherche. On peut préférer d'autres appellations, sans néanmoins pouvoir échapper à l'évaluation de l'informatisation des moyens de recherche.

– L'efficacité de cette informatisation ne doit pas faire oublier qu'elle présente le risque de favoriser une marginalisation de la réflexion sociopolitique.

BIBLIOGRAPHIE

Bachimont B., 2010. *Le sens de la technique : le numérique et le calcul*, Paris, Les Belles Lettres.

Bigot J.-É., 2018. *Instruments, pratiques et enjeux d'une recherche numériquement équipée en sciences humaines et sociales*. Thèse de doctorat, Université de Compiègne.

Boullier D., 2015. « Les sciences sociales face aux traces du big data. Société, opinion ou vibrations ? » *Revue française de science politique*, n° 65, pp. 805-828.

Broca S., 2016. « Épistémologie du code et imaginaire des "SHS 2.0" », *Variations*, n° 19. http://journals.openedition.org/variations/701. Dernière consultation le 12/03/2020.

Cocco Ch., Dessart G., Serbaeva O., Brandt P.-Y., Vinck D. et Darbellay F., 2018. « Potentialités et difficultés d'un projet en Humanités Numériques (DH) : confrontation aux outils et réorientations de recherche ». *Digital Humanities Quarterly*, 12(1). http://digitalhumanities.org/dhq/vol/12/1/000359/000359.html. Dernière consultation le 12/03/2020.

Compagno D., 2017. « Signifiant et significatif. Réflexions épistémologiques sur la sémiotique et l'analyse des données ». *Questions de communication*, n°31, pp. 49-70.

Juanals B. et Minel J.-L., 2017. *Enjeux numériques pour les médiations scientifiques et culturelles du passé*. Nanterre, Presses universitaires de Paris Nanterre.

Le Deuff O., 2018. *Les humanités digitales. Historique et développements*. Londres, ISTE Editions.

Longhi J., 2017. « Humanités, numérique : des corpus au sens, du sens aux corpus ». *Questions de communication*, n°31, pp. 7-18.

Manovich L., 2017. « Cultural Analytics, Social Computing and Digital Humanities ». Dans M. T. Schäfer et K. Van Es (eds). *The Datafied Society*. Amsterdam, Amsterdam University Press.

Mayaffre D., 2007. « Philologie et/ou herméneutique numérique : nouveaux concepts pour de nouvelles pratiques ? » dans F. Rastier et M. Ballabriga (dir.). *Corpus en Lettres et Sciences sociales. Des documents numériques à l'interprétation*. Toulouse, Presses universitaires de Toulouse.

Meunier J.-G., 2017. « Humanités Numériques et modélisation scientifique ». *Questions de communication*, n° 31, pp. 19-48.

Moretti F., 2016. « Abstraction, motifs récurrents et forme ». *Revue d'histoire littéraire de la France*, n° 116, pp. 521-532.

Mounier P., 2018. *Les Humanités Numériques. Une histoire critique*. Paris, Éditions de la Maison des sciences de l'homme.

Ouakrat A. et Mésangeau J., 2016. « Resocialiser les traces d'activités numériques : une proposition qualitative pour les Sic ». *Revue des sciences de l'information et de la communication*, n°8. http://journals.openedition.org/rfsic/1795. Dernière consultation le 12/03/2020.

Poibeau Th., 2014. « Le traitement automatique des langues pour les sciences sociales ». *Réseaux*, n° 188, pp. 25-51.

Puig V., 2017. « Web herméneutique et production de savoir : l'approche des Digital Studies ». *Mélanges de la Casa de Velázquez*, 47 (2), pp. 307-314.

Rastier F., 2008. « Sémantique du web *vs* semantic web ? Le problème de la pertinence ». *Syntaxe & Sémantique*, n° 9, pp. 15-36.

Schloen D. et Schloen S., 2014. « Beyond Gutenberg: Transcending the Document Paradigm in Digital Humanities ». *Digital Humanities Quarterly*, 8(4). http://digitalhumanities.org/dhq/vol/8/4/000196/000196.html. Dernière consultation le 12/03/2020.

Valluy J., version du 27/01/2020. *Transformations des États démocratiques industrialisés (TEDI)*. http://www.hnp.terra-hn-editions.org/TEDI/. Dernière consultation le 12/03/2020.

Venturini T., Jacomy M. et Jensen P., 2019. *What do we see when we look at networks*. https://arxiv.org/ftp/arxiv/papers/1905/1905.02202.pdf. Dernière consultation le 12/03/2020.

CHAPITRE II

Les Humanités Numériques, quelles définitions ?

Dominique Carré
Université Sorbonne Paris Nord, LabSic
EA 1803, MSH Paris Nord
Fanny Georges
Université Sorbonne Nouvelle, CIM EA 1484
Jérôme Valluy
Université Panthéon-Sorbonne (Paris 1), Costech EA 2223/
Université de Technologie de Compiègne

Résumé : Humanité et numérique : comment deux notions aussi éloignées ont-elles pu se relier et incarner une nouvelle orientation en matière de recherche et de système éducatif ? Après avoir ré-évalué dans le cadre d'une approche socio-historique les notions d'humanité et de numérique, il convient de s'interroger sur la manière dont advient un nouvel ordre, celui de la mathématisation de l'humain et de la société soulevant ainsi tout à la fois des enjeux à caractère épistémologique et idéologique. Ce processus se structure lors de la mise en oeuvre d'un triple processus : (1) structuration institutionnelle et labellisation ; (2) opportunisme disciplinaire ; et (3) injonction à dépasser les traditions de recherche. Les Humanités Numériques sont ainsi définies comme impliquant les conditions de production, de diffusion des connaissances et l'éthos du chercheur.

Mots-clés : Humanités Numériques, structuration institutionnelle, pratiques de recherche, mathématisation, données.

Abstract: Humanity and digital: how could two concepts so distant have been linked and embody a new direction in research and the education system? After having re-evaluated the notions of humanity and digital in the context of a socio-historical approach, it is appropriate to question the way in which a new order, that of the mathematization of the human and of the society thus raising both epistemological and ideological issues. This process is structured during the implementation of a triple process: (1) institutional structuring and labeling; (2) disciplinary opportunism; and (3) injunction to go beyond research traditions. Digital humanities are thus defined as involving the conditions of production, dissemination of knowledge and the researcher's ethos.

Keywords: digital humanities, institutional structuring, research practices, mathematization, data.

L'objectif de ce chapitre est de définir les Humanités Numériques. Mais avant de s'y pencher il semble nécessaire de comprendre comment deux termes aussi éloignés, pour ne pas dire opposés, ont pu se relier et devenir complémentaires en incarnant une nouvelle orientation d'étude ? Un nouveau champ de recherche ? Et peut-être, aux yeux de certains, une approche transdisciplinaire ?

« Humanités » et « numérique »

Que recouvre la notion d'« Humanités » ?

Les humanités découlent d'un courant de pensée, l'« humanisme » qui émerge à la Renaissance[1], bien que le terme « humanisme » ne date que de la seconde moitié du XIX^{e} siècle. L'Humanisme est un « mouvement d'esprit, une tendance générale, une conception du monde [...] Mouvement historique, force socio-culturelle » (Margolin, 1996, p. 728) qui renvoie à la perfection humaine, comme le note cet auteur, d'ordre éthique, esthétique et social. L'humanisme signifie que l'homme se construit de lui-même dans son accomplissement intellectuel, moral, religieux ; quant aux sciences à l'époque (pédagogie, philosophie, esthétique...), elles ont vocation à nous rendre plus humain. Érasme ne dit-il pas que l'humanité doit se dégager peu à peu de l'état de nature ? Pour mener à bien ce processus, il attachera un soin particulier à l'action pédagogique. N'oublions pas qu'en France les humanités sont liées à la culture humaniste et au système éducatif.

Les humanités dans notre pays recouvrent ce qui est le plus souvent appelé les « humanités classiques » : les lettres, l'esthétique, la rhétorique. *A contrario* l'approche anglo-saxonne prend en compte un spectre beaucoup large allant des arts, en passant par les lettres, la philosophie jusqu'aux sciences humaines et même sociales. Durant

1. Nous retenons cette temporalité même si certains auteurs comme Marie-Madeleine Compère et André Chervel (1997) indiquent que les humanités remontent aux « arts libéraux » antiques.

la seconde partie du XX[e] siècle les disciplines comme la philosophie, l'histoire, la littérature, l'anthropologie etc. se sont retrouvées dévaluées. Marginalisées, peu attractives « certains pensent pouvoir, malgré tout, sauver une tradition érudite et multiséculaire en la refondant dans le grand creuset numérique » (Mounier, 2010, pp. 10-11). Ainsi les humanités tombées en désuétude, longtemps suspectes et tenues à distance semblent refaire depuis une ou deux décennies un retour en force (Denizot, 2015). Depuis quelque temps les universités développent à nouveau des cursus d'humanités. On assiste progressivement pour des raisons diverses à l'« autonomisation d'un champ » comme l'indique fort justement Denizot qui regroupe selon les universités des disciplines différentes en sciences humaines et sociales comprenant parfois les sciences de l'information de la communication (Sic). Même l'École Polytechnique possède un département Humanités. Notons que le Ministère de l'Éducation Nationale vise à installer aussi dans les programmes de Lycée à côté des « Humanités scientifiques », des « Humanités » mais cette fois « numériques ».

Que recouvre la notion de « numérique » ?

Le mot numérique désigne classiquement, notamment dans les dictionnaires, ce qui se rapporte aux nombres, ce qui est représenté par un ou plusieurs nombres, et renvoie alors à la discipline des mathématiques. La dimension informatique du concept n'apparaît dans les dictionnaires que récemment et sous des formulations encore très limitatives aujourd'hui, par comparaison avec le sens observable donné à ce mot dans la presse et dans le langage courant. Ce sens désigne plusieurs aspects d'une transformation globale qui s'opère principalement au début du XXI[e] siècle dans les sociétés développées : la généralisation des usages sociaux de technologies informatiques, servies par des innovations en matière d'ordinateurs et de réseaux, permettant à des personnes non spécialisées de les utiliser intuitivement – sans maîtriser le code informatique – pour communiquer. Dans cet usage récent, le mot numérique évoque à la fois une continuité entre informatique et numérique mais plus encore et surtout une distance tant, en 2019, la compétence informatique ne fonde plus l'usage, par de vastes populations, des technologies concernées. L'élévation des taux d'équipement en ordinateurs portables, tablettes tactiles et *smartphones*, dans les pays occidentaux les plus riches, et la production industrielle d'applications destinées à des utilisateurs néophytes dans tous les domaines et aspects de la vie

en société, modifie en une ou deux décennies, le sens même du mot « numérique ». Il désigne alors, sinon une société « nouvelle » que serait une « société numérique », du moins une dimension nouvelle de la société contemporaine en tant qu'elle se caractérise par cette généralisation des usages sociaux de technologies d'information et de communication autant que par les conséquences globales de cette généralisation sur la configuration de la société elle-même, sa culture notamment et son système économique ; comme l'observe Milad Doueihi « Le contexte change et avec ce changement advient un nouvel ordre, celui des données et de leurs interprétations par les divers modèles de classement algorithmique. On est tenté de voir dans ce tournant une variation numérique sur le « fait social total » introduit jadis par Marcel Mauss. » (Doueihi, 2011, p. 12).

Humanités Numériques ou digitales ?

L'expression humanités « numériques » ou « digitales » fait partie des nombreuses expressions[2] qui sont apparues dans le monde académique pour désigner les approches en sciences humaines et sociales qui s'appuient sur les technologies numériques comme méthode de traitement des données, qu'elles constituent ou non un terrain d'étude des pratiques sociales ou de la culture numérique.

Utilisée au départ plutôt pour désigner les différentes approches bibliométriques et en traitement automatique des langues et valoriser les grandes entreprises de numérisation et d'exploitation de grands corpus littéraires et la constitution d'archives numériques dans les bibliothèques, le débat semble s'être fixé sur cette expression, qui selon les auteurs qui y recourent, devrait conserver une nébulosité pour accueillir toutes les initiatives qui souhaitent s'y inscrire[3].

Lorsqu'au substantif « humanités » est accolé le qualificatif « numérique », les « Humanités Numériques » (ou plus exactement les humanités dites numériques) sont devenues un domaine de recherche et d'enseignement multiforme au croisement de l'informatique des arts, lettres, sciences humaines et sociales. Elles interrogent les conditions de production et de diffusion des connaissances. Notons

2. *Literary and Linguistic Computing*, Humanities Computing, *Digital Humanities*, informatique littéraire, *informatica umanistica*, *digitale Geisteswissenschaften* (Berra 2015) ; évolution du domaine de l'édition numérique « *scholarly communication* » *Software Studies*, Media Studies, Digital Studies (Gras, 2015). Selon Bouzidi et Boulesnane (2017) l'origine du terme *Digital Humanities* est attribuée à Schreibmann, Siemens et Unsworth (2004) avec l'édition d'un ouvrage de référence qui avait pour intitulé *A Companion to Digital Humanities*.
3. Svensson, 2012 Cité par Plantin, 2014, p. 47.

qu'elles ne sont pas reconnues jusqu'à présent par le Conseil National des Universités (CNU).

Les Humanités Numériques ne résultent pas d'une génération spontanée. Historiquement, on peut estimer qu'elles sont le fruit de ce qui a été appelé après la seconde guerre mondiale les « *Humanities computing* » (sciences humaines assistées par ordinateur). Il est possible d'identifier d'une manière non exhaustive deux précurseurs des « Humanités Numériques ». D'une part, le père jésuite Roberto Busa, qui dès 1946 aurait commandé la transformation des textes de *La Somme théologique* de *Saint Thomas d'Aquin* en fiches perforées à Thomas J. Watson, fondateur et « *patron* » d'IBM en 1949 (Le Deuff, 2015, Berra 2015, Doueihi 2015) a eu recours à l'ordinateur (IBM) pour faciliter son étude sur *de Thomas d'Aquin*. Cette fable à la coloration quasi-mythique[4] ou héroïque[5] diffuse subrepticement une morale que les technologies numériques, émanations du capitalisme, placées sous l'égide de l'impérialisme européen, se transfigureraient en des technologies du texte et des savoirs. Des valeurs charitables seraient ainsi permises par le projet des Humanités Numériques à travers cette ascendance : « sauver des traditions d'étude menacées par le discrédit et la pénurie » (Berra, 2015).

D'autre part, Jean-Claude Gardin qui fonde en France, en 1957 au CNRS le Centre mécanographique de documentation archéologique (CMDA) qui deviendra en 1960 le Centre d'analyse documentaire pour l'archéologie (CADA). Comme le précise Roger Bautier (2017), les travaux de Gardin vont promouvoir un programme logistique applicable aux sciences humaines et sociales (tout particulièrement à l'archéologie). L'objectif est d'améliorer fortement la valeur scientifique de ces sciences. Quant au terme « Humanités Numériques » il apparaît la première fois en 2004 dans une publication collective : *A Companion to Digital Humanities* (Schreibman, Siemens, Unsworth, 2004). Il est nécessaire à ce stade d'apporter une précision. *Digital* est en règle générale traduit en français par numérique. Le syntagme des « études digitales », propulsé dans le champ francophone notamment par l'ouvrage collectif dirigé par Bernard Stiegler en mars 2014,

4. « Au commencement étaient les œuvres de Thomas d'Aquin, puis vers 1949 vint Roberto Busa, son commentateur jésuite et notre père fondateur, qui alla trouver le patron de l'International Business Machines Corporation (IBM) et mit en fiches perforées l'Opus Thomisticum pour en élaborer la concordance. Ainsi débute le récit canonique proposé dans un companion book cent billets de blog et dans les pages de Wikipédia. » (Berra, 2015). Voir également Doueihi (2015), Plantin (2014).

5. Bénel (2014), parle d'une héroïsation du père Busa sur la page Wikipédia américaine.

apparaît comme l'expression d'une première divergence qui s'affirmera deux ans plus tard au sein des sciences de l'information et de la communication. Ce livre ouvre un espace de débat entre Humanités Numériques et *Digital Studies*, ici entre ceux qui pensent qu'il faut traiter surtout les effets du numérique sur l'homme et la société, et ceux qui préfèrent s'occuper des méthodes et des dispositifs que le numérique met à la disposition des humanités et/ou des sciences humaines et sociales (Stiegler, 2014).

Mais pour Olivier Le Deuff (2015) il est plus pertinent de retenir le qualificatif « digital » que « numérique » pour une autre raison. Si l'on suit le raisonnement de cet auteur cela permettrait de mieux inscrire les humanités digitales dans une histoire longue qui précède l'émergence du numérique et de l'informatique[6]. En effet, celui-ci considère que l'apparition de l'index marque l'émergence des humanités digitales bien avant l'arrivée de l'informatique. Que l'on retienne les expressions « Humanités Numériques » ou « humanités digitales », elles se distinguent de l'« humanisme numérique » proposé par Milad Doueihi (2011) qui s'inscrit dans une dimension ontologique et philosophique. Façonnant l'humain, le numérique crée de nouvelles humanités, comme le résume Vitali-Rosati (2011) ; et de l'« Humanité Numérique » (au singulier) qui caractérise, le plus souvent, l'impact des traitements numérisés sur l'humain.

Rappelons à ce stade que les travaux du symposium SFSIC-CPdirsic et de l'atelier consacré aux Humanités Numériques ont permis de mettre en évidence trois approches complémentaires des Humanités Numériques :

— La première est centrée sur la collecte, la gestion, la diffusion des données numériques au service des Sciences Humaines et Sociales (SHS) ;

— La deuxième renvoie aux travaux des informaticiens en matière d'Intelligence Artificielle (IA) et de traitement des corpus documentaires au service des organisations (*Big Data*, par exemple) ;

— La troisième, quant à elle s'inscrit dans une archéologie des médias, anciens ou numériques et des dispositifs matériels d'information et

6. Cet auteur rappelle que cette histoire est celle des pratiques lettrées, mais surtout l'histoire des outils de classement et d'organisation de la pensée et des connaissances. En fait celui-ci veut nous faire comprendre pour reprendre ses termes que « les *Humanités Digitales* s'inscrivent dans une lignée d'une lecture organisée qui sort de la linéarité du texte et du document ».

de communication (non loin des travaux inspirés par la médiologie francophone, comme dans la médiologie des traces numériques).

Les Humanités Numériques en Sciences de l'Information et de la Communication

L'état de la recherche portant sur les Humanités Numériques et digitales dans les revues francophones reconnues en Sciences de l'Information et de la Communication laisse apparaître plusieurs traits saillants définissant ce nouveau champ que nous allons présenter dans cette partie : la dimension institutionnelle des Humanités Numériques ou digitales, la numérisation des outils de la recherche, le dépassement de la dichotomie sciences dures-sciences molles par la quantification, la mutation du *modus operandi* et de l'éthos du chercheur.

Les Humanités Numériques ou digitales sont un label inscrit dans un processus de structuration institutionnelle. De nouveaux collectifs pluridisciplinaires[7] apparaissent depuis le début des années 2000. Ce nouveau label peut être perçu comme une injonction associée au besoin de financement et de revalorisation, voire de réinvention, de domaines disciplinaires peu familiers de ces procédures. Adhérant à une pseudo-« culture par projet » issue du monde de l'entreprenariat (Bénel, 2014, p. 104) imprégnée d'une certaine conception de la « culture en réseau » (Soulier, 2014, p. 25), les Humanités Numériques attirent tout autant qu'elles agacent, au point que les frictions, tensions, opportunismes (Doueihi, 2015, p. 108), et arrogances disciplinaires (Berra, 2015, p. 619) pourraient en sembler constitutives. De fait, elles sont animées par une véritable injonction paradoxale : alors qu'elles s'appuient et construisent le tissu institutionnel, les Humanités Numériques semblent se stabiliser dans leurs difficultés à trouver une forme, comme « zone expérimentale permettant la mise en place de recherches orientées informations et données, distribuées, collaboratives et pluri-disciplinaires » (Borgman, 2009, cité par Plantin, 2014, p. 48).

La numérisation des outils de la recherche implique tout le parcours du chercheur et de son activité de recherche, de la formation du chercheur jusqu'à la diffusion de la production scientifique :

7. Bouzidi et Boulesnane (2017, pp. 23-24) dressent une liste des réseaux internationaux dont CenterNet animé par le Canadian Institute for Research in Computing and the Arts, le Center for Digital Research in the Humanities aux États-Unis, le *Department of Digital Humanities*, en Angleterre, le *Digital Humanities Hub* en Australie et dénombrent près de 27 conférences annuelles.

environnement personnel de travail, mode de médiation des résultats, travail de terrain, recherches historiographiques, archives de la recherche. La maîtrise de ces outils et la construction d'un protocole de recherche interdisciplinaire ou pluridisciplinaire impliquent que la « zone d'expérimentation » soit aussi une *zone d'échange* dans laquelle les collaborateurs puissent partager librement l'accès aux données (Plantin, 2014, p. 46) par des outils numériques permettant tour à tour de numériser, extraire les données, les exploiter, échanger, voire rédiger et diffuser les résultats.

Les projets en Humanités Numériques auraient vocation à dépasser la traditionnelle dichotomie entre les deux traditions de recherche traditionnellement séparées de l'*aesthesis* des humanités (connaissance subjective), et la *mathesis* des sciences exactes (l'objectivité) (Plantin, 2014, p. 46). Le premier défi à relever est le dépassement de la dichotomie quantitative-qualitative.

> « La méthode qualitative fournit un raisonnement sur les actions humaines, basé sur des démarches ouvertes et participatives. Tandis que la méthode quantitative permet d'appuyer un raisonnement, de vérifier des faits ou de définir des variables à travers des méthodes computationnelles et statistiques, entre autres. Les actions sont souvent simplifiées par l'usage d'outils numériques avancés qui favorisent l'automatisation des tâches. » (Bouzidi et Boulesnane, 2017, p. 22.)

Sur le chemin de cette association nouvelle, l'importation des méthodes quantifiées en sciences humaines a pu donner lieu à des débordements[8] ou à des figures d'autorité déplacées du *nombre*, agitant dans les esprits une « notion réductrice du progrès[9], qui ont peu donné de gages de crédibilité au nouveau champ. Cependant, la quantification de la méthode qualitative ou le traitement de corpora de masses de données en recourant à l'information et à l'intelligence artificielle ont construit progressivement un cadre méthodologique

8. « Les graphes sont faciles à produire mais pas également lisibles, ni utiles. Jockers présente un cas limite avec ses chevelures 10 de milliers de livres reliés par distance thématique. En jouant sur les tailles, les couleurs, les transparences, on peut produire des taches évolutives et psychédéliques, certainement propices à l'ouverture des portes de la perception, mais dont la valeur démonstrative est moins évidente. Les graphes n'ont pas vocation à être spectaculaires : ils devraient devenir un outil d'analyse aussi naturel qu'un graphique "camembert" et être regardés avec la même distance critique. » (Glorieux, 2015, p. 661).

9. « Les discours portant sur « le » numérique ravivent, consciemment ou non, une notion réductrice du progrès. Positivisme, scientisme, désir indu de mimer les sciences » (Berra, 2015, p. 617-618).

plus solide. Des questionnements et perspectives nouvelles en littérature, philologie, sociologie, inspirées de la biologie – telle l'approche phylogénétique, les approches cartographiques ou archéologiques. Parmi les nouveaux outils conceptuels et méthodologiques, figurent la modélisation, l'usage d'outils de TEI, de réseaux sociaux (NodeXL), de liens hypertextes (Navicrawler), les approches par graphes. La question de l'archivage est structurante pour les Humanités Numériques, mais des malentendus planent encore sur la donnée, centrale pour définir ce qu'il faut archiver. La donnée scientifique, multipliée et révélée sous de nouveaux aspects, est réexaminée et réinterrogée en profondeur. Elle semble renvoyer chaque discipline à son propre regard. Les littéraires voient en la donnée du texte numérisé, les sociologues des propos numérisés, la sémiologie des traces. Pour les chercheurs qui travaillent sur la socialisation médiée par le numérique, le numérique implique une ré-interrogation radicale de ce qui est donné à voir sur les nouveaux espaces numériques alors considérés comme espaces de conduite de la socialisation ordinaire.

L'introduction des outils numériques dans le *modus operandi* du chercheur implique un bouleversement des manières de conduire la recherche, mais aussi de la formation des chercheurs, de la littératie numérique, qui impliquent elles-mêmes un bouleversement de l'autorité scientifique[10] et de l'éthos du chercheur qui frisent une forme de dissidence conforme à l'humanisme ancien. Sous leur aspect le plus radical, en opposition aux valeurs traditionnelles de la recherche, les Humanités Numériques peuvent prendre le visage de l'enfant terrible de l'institution : sous son étendard certains prônent le refus de la hiérarchie (au profil d'une hétérarchie, voir Soulier, 2014, pp. 25-26) et le refus du cloisonnement disciplinaire. En leur place sont promues les valeurs d'ouverture, d'entraide, de bienveillance, d'esprit d'initiative, d'inclusion, de transparence, de collaboration, de respect mutuel entre individus, entre communautés, institutions et disciplines.

L'obtention du label institutionnel, l'impératif de collaboration, l'adoption de méthodes impliquant le numérique, modifient en profondeur les pratiques de recherche et l'éthos du chercheur. Par-delà les enjeux stratégiques de l'obtention du label, ces nouvelles configurations mettent en évidence de nouveaux questionnements

10. « Valoriser la dimension active, au détriment du pôle contemplatif de la connaissance, c'est aussi faire un pied de nez à la recherche solitaire, à la spécialisation académique et à l'autorité scientifique. » Soulier, 2014, p. 23.

épistémologiques sur la nature de ce nouveau savoir. On observe en particulier un « déplacement que subit la lecture érudite et savante vers un modèle épistémologique fondé sur une conception quantitative de la preuve » (Doueihi, 2015). À l'horizon de ces nouvelles perspectives de recherche, d'aucuns défendent l'idée que les Humanités Numériques devraient mettre en œuvre une manière méthodologique commune de penser et de faire de la recherche (Bouzidi et Boulesnane, 2017, p. 22), un « bien commun méthodologique (et codique) » (Soulier, 2014, p. 25).

Vers une définition large des Humanités Numériques ?

Sur le plan intellectuel, scientifique, philosophique ou esthétique, le label « Humanités Numériques » demeure indéfini en ce qui concerne ses contenus et limites, comme domaine de recherche et d'enseignement, autant que peu structuré dans ses orientations théoriques ou philosophiques[11]. L'étude comparée de corpus bibliographiques anglophones et francophones réalisée par Églantine Schmitt fait apparaître en outre une forte différenciation des définitions implicites, presque culturelles, du domaine émergent des Humanités Numériques : « les publications francophones occupent donc un champ thématique et disciplinaire beaucoup plus restreint que leurs pendants anglophones. Contrairement à ce qu'ambitionne le manifeste francophone des Humanités Numériques, l'inclusion des sciences humaines et sociales et l'interdisciplinarité avec l'informatique ne sont pas perceptibles, si l'on en croit l'analyse présentée ici, dans les travaux effectivement publiés autour des *digital humanities* » (Schmitt 2015)[12]. Cette difficulté qui paraît liée à une culture académique francophone ancienne relative à la définition restrictive des humanités (classiques) pourrait inciter à élargir la définition des Humanités Numériques afin de satisfaire à cette ambition d'inclusion des sciences humaines et sociales.

En effet, malgré (ou à cause de) leur relative indéfinition, les « Humanités Numériques » forment le premier label ou intitulé assez largement utilisé pour fédérer et dynamiser les efforts de prise de conscience et d'adaptation au sein de l'ensemble des sciences humaines. Deux disciplines occupent des positions spécifiques : les sciences de l'information et de la communication (Sic) et les sciences et ingénieries de l'informatique (de plus en plus souvent labellisées

11. Cette troisième partie reprend des parties ou s'inspire d'un texte collectif coproduit en 2017 : « Humanités Numériques plurielles » : http://www.reseau-terra.eu/spip.php?rubrique304.
12. L'auteur évoque dans son texte le manifeste de 2010 (Mounier, 2010).

« Sciences et technologies de l'information et de la communication » STIC) sont particulièrement actives et productives sur le domaine, avec ou sans l'usage du label « Humanités Numériques ». Dans toutes les autres disciplines de sciences humaines, ce label est le seul, aujourd'hui, permettant de désigner de façon transversale à ces disciplines la fraction, marginale en chacune d'elles, d'objets d'étude, de travaux, d'auteurs, de débats se rapportant aux dimensions numériques de chacune des disciplines et de leurs domaines d'étude et d'enseignement respectifs. C'est probablement l'une des raisons qui a conduit à insérer le label « Humanités Numériques » en février 2014 dans la nomenclature nationale des formations universitaires de niveau Master[13]. En ce sens, ce label semble exprimer un appel aux arts, lettres et sciences humaines et sociales à réfléchir à leur agenda intellectuel et scientifique, à leurs méthodes de travail et à leur place dans une société du XXI^{e} siècle en cours de transformation rapide sous l'effet de la généralisation des usages sociaux de technologies numériques. Ces constats sur le champ des sciences humaines incitent aussi à soutenir la dimension potentiellement fédératrice du label mais, nécessairement, avec une définition large du domaine de recherche et d'enseignement.

Le « Manifeste des *Digital Humanities* » publié en 2010 plus tôt dans le *Journal des anthropologues* retenait une définition technique voire techniciste du domaine : « L'utilisation de l'informatique en sciences humaines et sociales est pratiquée depuis maintenant plus de quarante ans. Plusieurs voies ont été explorées au cours de cette déjà assez longue histoire. La plus récente, qui prend le nom de *digital humanities*, désigne une intégration intense et à plusieurs niveaux des technologies numériques dans tous les processus de recherche, depuis la collecte de données jusqu'à la publication. » (Mounier, 2010). L'année 2010 est aussi l'année de création du SYNTEC-Numérique (Syndicat professionnel des Entreprises de Services du Numérique, du Conseil en Technologie et des Éditeurs de Logiciels) marquant la reconnaissance dans ces entreprises des nouveaux métiers du numérique avec notamment un intérêt marqué pour les compétences relatives aux traitements automatisés de corpus non structurés de données issues, en particulier, des réseaux sociaux et du web. La même perspective définitionnelle se trouve imposée dans l'article « humanités_numériques » de la Wikipédia francophone notamment

13. Arrêté du 4 février 2014 fixant la nomenclature des mentions du diplôme national de master, JORF n° 0035 du 11 février 2014 p. 2414 texte n° 21 : https://www.legifrance.gouv.fr/eli/arrete/2014/2/4/ESRS1402516A/jo/texte.

à partir de 2013 : l'étude systématique de son historique sur la période 2012/2016 montre que la définition du label est un enjeu qui donne lieu à des stratégies d'affichage, voire de publicité et de luttes entre points de vue distincts. L'article est sévèrement gardé voire capté : sur les quatre années d'évolution, on peut estimer à une dizaine seulement, le nombre de personnes ayant contribué à la plus grosse proportion d'écritures sur cette page. De cette définition ne sont pas absents les espoirs scientistes d'une mathématisation de l'humain et de la société. Ainsi Serge Abiteboul et Florence Hachez-Leroy indiquent dans leur article « Humanités Numériques » de l'Encyclopédie de l'humanisme méditerranéen : « Les humanités sont donc en train de se réinventer par l'informatique et de se rapprocher des autres sciences, nous semble-t-il. Ces convergences sont si fortes que plutôt que de parler des Humanités Numériques, peut-être aurait-il fallu discourir de "sciences numériques" en général. » (Abiteboul, Hachez-Leroy 2014). Néanmoins ces approches des Humanités Numériques ne sont pas apparues dans l'espace francophone sans générer de débats scientifiques voire de controverses ces trois dernières années notamment (Bonaccorsi, Carayol, Domenget, 2016 ; Granjon, 2016) : leur sont reproché une sous-estimation de la dimension sociale préalable à toute conception technologique, une faible capacité à rendre compte des différenciations sectorielles et disciplinaires du numérique (de ces technologies elles-mêmes et leurs usages variables selon les secteurs socioprofessionnels), une tendance à propulser ou relayer des illusions scientistes quant au rôle des statistiques dans la connaissance des faits sociaux, un trop faible examen critique des conditions socio-économiques de production des données massives…

Ces discussions scientifiques des dernières années incitent également à adopter une définition élargie des Humanités Numériques qui ne serait que pour préserver un espace commun de dialogue voire de controverses utiles entre les différents points de vue qui sont apparus autour de ce domaine. Dans cette perspective, les sciences de l'information et de la communication apparaissent prédisposées à favoriser cette définition large et fédérative. En raison de leurs avancées dans ce domaine et de leur pluralisme intellectuel, les Sic peuvent aider à définir les Humanités Numériques autour de six axes favorables à leur pluralisme intellectuel[14] :

14. Ces propositions reprennent la formulation résumée du document collectif précité ; Jérôme Valluy, « Bénéficier des critiques et controverses relatives aux Humanités Numériques pour en construire une définition pluraliste », *Revue française des sciences de l'information et de la communication*, 10 | 2017 : http://journals.openedition.org/rfsic/2656.

1– Inscrire les H. N. dans la modernité des « humanités », telles qu'elles se présentent au XXI[e] siècle, non pas limitées aux humanités classiques (langues anciennes, lettres, arts, philologie et historiographies connexes) ni focalisées seulement sur les archives, le patrimoine, l'histoire, les lettres…, mais ouvertes aux enjeux épistémologiques et méthodologiques ainsi qu'aux apports de connaissances des SHS dans leur diversité.

2– Placer au premier plan de l'agenda scientifique l'étude des transformations de la société sous l'effet des technologies numériques (*i.e* : que nous apprennent les humanités modernes sur le tournant numérique de la société ?) sans focalisation sur les seuls outils, instruments et potentialités qu'offrent les nouvelles technologies apparaissant chaque année (*i.e* : que peut-on faire avec les outils numériques dans les humanités modernes ?).

3– Assumer la charge de travail en recherche pour identifier et analyser les spécificités du numérique dans chaque secteur de la société (secteurs d'activités professionnelles ou citoyennes et secteurs disciplinaires de la connaissance scientifique) contre les simplifications globalisantes dans l'interprétation de la dimension numérique du monde et pour améliorer à terme, par cumul de recherches spécialisées, la compréhension de cette dimension.

4– Assumer aussi la charge de travail d'étude des différenciations de conceptions et d'usages du numérique d'un champ linguistique à l'autre, loin des visions stratosphériques du numérique planétaire par examen limité aux tendances technologiques internationales, et examiner l'hypothèse comparative de différenciation des champs linguistiques et culturels du numérique.

5– Retrouver dans les H. N. le pluralisme axiologique, théorique et paradigmatique qui a fait la richesse des humanités modernes en favorisant l'esprit critique de chaque point de vue vis-à-vis des autres, donc la controverse, vertueuse, par ses effets heuristiques. Cela nécessite de construire l'indépendance (individuelle et collective) intellectuelle des chercheurs face aux idéologies dominantes, injonctions politico-médiatiques et cadrages technocratiques.

6– Favoriser la diversité des formes techniques, cadres organisationnels et modèles économiques de l'édition numérique, constitue une condition essentielle pour faire prospérer ce pluralisme ; cela implique d'expérimenter des formats éditoriaux numériques, des dispositifs de validation scientifique et des modalités pratiques d'éditorialisation qui articulent dans les H. N. deux valeurs connexes : le libre accès aux savoirs et l'indépendance des auteurs dans leur diversité.

Proposition

Les Sic en raison de leur ancienneté et avance dans les recherches sur la dimension numérique de la société, autant que leur pluralisme intellectuel favorisé par une orientation interdisciplinaire ancienne devraient adopter une définition large et fédérative des Humanités Numériques. Cette définition consiste à articuler étroitement dans cette spécialité émergente les recherches sur l'apport des outils et données numériques aux sciences humaines et les recherches sur les contextes sociétaux d'émergences et usages de ces outils, ainsi que sur les perceptions sociales de ce que sont des « données », en considérant les différences entre espaces linguistiques de communications numériques, entre contextes nationaux et entre secteurs socioprofessionnels. Cette définition large aidera à dépasser la dichotomie des approches quantitatives/qualitatives pour mieux articuler des méthodes dont la complémentarité est souvent indispensable. Elle favorisera les dialogues interculturels relatifs aux Humanités Numériques notamment entre les champs académiques anglophones et francophones pour lesquels la notion d'« humanités » a des significations différentes. Cette définition large valorise le pluralisme axiologique, théorique et paradigmatique qui fait la richesse des sciences humaines en favorisant l'esprit critique, donc la controverse, vertueuse, par ses effets heuristiques. Cette conception ouverte des Humanités Numériques soutient la diversité des formes techniques, cadres organisationnels et modèles économiques de l'édition numérique, et valorise l'expérimentation de nouveaux formats éditoriaux, dispositifs de validation scientifique et modalités pratiques d'éditorialisation en articulant deux valeurs connexes : le libre accès aux savoirs et l'indépendance des auteurs dans leur diversité.

POSITIONNEMENT

Notre proposition centrale consiste à adopter en Sic une définition large et fédérative des Humanités Numériques afin de favoriser les dialogues mutuellement enrichissants entre cette spécialité émergente et les autres spécialités de la discipline d'une part et de favoriser les compréhensions interculturelles sur ce domaine entre champs académiques anglophones et francophones d'autre part. Cette définition large devrait articuler étroitement les recherches sur l'apport des outils et données numériques aux sciences humaines et les recherches sur les contextes sociétaux d'émergences et usages de ces outils, ainsi que sur les perceptions sociales de ce que sont des « données » en tenant compte des différences entre champs linguistiques, pays et secteurs socioprofessionnels.

Cette définition large devrait inclure les recherches sur la diversité des formes techniques, cadres organisationnels et modèles économiques d'édition numérique, et sur l'expérimentation de nouveaux formats éditoriaux, dispositifs de validation scientifique et modalités pratiques d'éditorialisation en articulant deux valeurs connexes : le libre accès aux savoirs et l'indépendance des auteurs dans leur diversité.

BIBLIOGRAPHIE

Abiteboul S. et Hachez-Leroy F., 2014. « Humanités Numériques » in Touati H. (dir.), *Encyclopédie de l'humanisme méditerranéen*. https://www.encyclopedie-humanisme.com/?Humanites-numeriques. Dernière consultation le 12/03/2020.

Bautier R., 2017. « Quelle place pour le logicisme dans l'organisation et la communication des connaissances », in Mustafa El Hadi W. (dir.). *Fondements épistémologiques et théoriques de la science de l'information-documentation,* London, Iste Editions.

Bénel A., 2014. « *Quelle interdisciplinarité pour les* Humanités Numériques ? » *Les Cahiers du numérique,* n° 4, vol. 10, pp. 103-132. https://www.cairn.info/revue-les-cahiers-du-numerique-2014-4-page-103.htm. Dernière consultation le 12/03/2020.

Berra A., 2015. « Pour une histoire des Humanités Numériques », *Critique*, 8, 819-820, pp. 613-626.

Bonaccorsi J., Carayol V. et Domenget J.-C. (dir.), 2016. « Humanités Numériques et Sciences de l'Information et de la communication », *Revue française des sciences de l'information et de la communication*, n° 8. https://doi.org/10.4000/rfsic.1778. Dernière consultation le 12/03/2020.

Bouzidi L. et Boulesnane S., 2017. « Les Humanités Numériques. L'évolution des usages et des pratiques. » *Les Cahiers du numérique*, 3, pp 19-38.

Compère M.-M. et Chervel A., 1997. « Les humanités dans l'histoire de l'enseignement français », *Histoire de l'éducation*, n° 74, pp. 56-38.

Denizot N., 2015. « Les humanités, la culture humaniste et la culture scolaire », *Tréma*, n° 43, pp. 42-51.

Doueihi, M., 2011. *Pour un humanisme numérique*, Paris, Seuil.

Doueihi M., 2015. « Quelles Humanités Numériques ? », *Critique*, n° 819-820, pp. 704-711.

Glorieux F., 2015. « Bibliothèques informatisées et lectures distantes », *Critique*, vol. 8, n° 819-820, pp. 653- 666.

Granjon F. (dir.), 2016. « Critique des Humanités Numériques », *Variations*, n° 19.

Le Deuff O., 2015. *Les humanités digitales précèdent-elle le numérique* ? Hal, 27 octobre : https://archivesic.ccsd.cnrs.fr/sic_01220978/document. Dernière consultation le 12/03/2020.

Margolin J.-C., 1996. « Humanisme », *Encyclopædia Universalis*, pp. 727-729.

Mounier P., 2010. « Manifeste des Digital Humanities », *Journal des anthropologues*, 122-123 (3), pp. 447-452.

Plantin J.-C., 2014. « Les digital humanities. Accomplissements et défis pour un agencement post-disciplinaire », *Les Cahiers du numérique*, vol. 4, n° 10, pp. 41-62.

Schmitt E., 2015. « La structuration disciplinaire et thématique des Humanités Numériques », Colloque *ISKO-France* : *Systèmes d'organisation de connaissances et Humanités Numériques*. Strasbourg.

Schreibman S., Siemens R. et Unsworth J. (dir.), 2004. *A Companion to Digital Humanities*, Hoboken (New Jersey), Wiley Blackwell.

Stiegler B. (dir.), 2014. *Digital Studies* : *organologie des savoirs et technologies de la connaissance*, Limoges, FYP Editions, 2014.

Soulier E., « Introduction. Les Humanités Numériques sont-elles des agencements ? », *Les Cahiers du numérique*, n° 10, pp. 9-40.

Vitali-Rosati M., 2011. « Une philosophie du numérique. Lecture de *Pour un humanisme numérique de Milad Doueihi* ». Montréal, Université de Montréal. https://papyrus.bib.umontreal.ca/xmlui/handle/1866/12996. Dernière consultation le 12/03/2020.

CHAPITRE III

Les Humanités Numériques, depuis les Sic

Chérifa Boukacem-Zeghmouri
Université Lyon 1
Françoise Paquienséguy
Sciences Po Lyon

Résumé : Confrontées aux technologies et à leurs significations sociales, à la nécessaire évolution des modes de traitement de l'information ou de la donnée, les Sciences de l'Information et de la Communication ont porté des préoccupations et des recherches qui participent à la genèse des Humanités Numériques. Pour en apprécier l'importance et la spécificité, ce texte aborde les principaux jalons qui ont amené les Sic à se confronter à des problématiques qui émargent aujourd'hui autour des Humanités Numériques.

Mots-clés : numérique, interdisciplinarité, *data*, document, médias, société.

Abstract: Information and Communication Sciences as a discipline has brought concerns and relevant research that do contribute to the genesis of Digital Humanities. This is due to the fact that the discipline has always been confronted with technologies, their social meanings, and the necessary evolution of information or data processing methods.The chapter adresses and highlights the main milestones that let Information and Communication Sciences related to issues considered as part of what is currently defined as Digital Humanities. By doing so, the chapter provides an original frame to start a discussion on the relationship between Information and Communication Sciences and Digital Humanities.

Keywords: Digital, Interdisciplinarity, Data, Document, Media, Society.

Les Sciences de l'Information et de la Communication (Sic) constituent un lien étroit et fort entre les évolutions majeures de la société depuis leur naissance et ce lien constitue lui-même leur raison d'être. En effet, dans les années 1970, les premières réflexions scientifiques qui ont posé les fondamentaux et les fondations de la discipline ont avancé que l'information et sa propagation ou sa transmission appellent leurs propres appareils et paradigmes théoriques nécessaires à des analyses et travaux longitudinaux, complémentaires des sondages quantitatifs qui abondent, l'informatique soutenant la statistique. La posture critique s'est construite progressivement pour saisir, à défaut de pouvoir juguler, l'utopie portée par la communication (Mac Luhan, Wiener). D'autant plus que la communication s'est trouvée-là renforcée par les premiers déploiements des technologies de la communication (au premier rang desquelles figure le téléphone ou le fax), jusqu'à conquérir la société tout entière (Miège, 1980). Ce sont ces phénomènes sociaux et techniques qui vont donner matière aux Sciences de l'Information et de la Communication et leur permettre de se structurer autour de leur objectif premier, étudier « la diversité des champs de l'information-communication en s'attachant notamment aux phénomènes de médiation, de conception, de production et de réception, aux représentations, à l'appropriation des dispositifs sociotechniques et à l'innovation » (définition CNU septembre 2019).

Sans refaire ici l'histoire de la discipline qui a déjà été bien instruite d'une part, et de façon contrastée de l'autre, rappelons simplement que les mutations et développements de la société française depuis les années 1970 ont entraîné une expansion manifeste des termes information et communication, mais aussi des stratégies, des discours et des usages relatifs à cet ensemble qui a été alors dénommé, sans beaucoup de précision à vrai dire, les NTIC (Nouvelles Technologies de l'Information et de la Communication), infiltrées dans des activités de plus en plus nombreuses et variées des sphères professionnelles

et privée. La communication est le propre de toute activité sociale et s'articule autour des informations qui en font le substrat. C'est pourquoi les Sciences de l'Information et de la Communication se sont constituées à partir d'autres disciplines déjà instituées, et se revendiquent, en France, porteuses d'une posture particulière qui lie les deux vocables, s'appuie sur un pluriel et la déclare interdisciplinaire. Cette singularité, française, en fait l'appareil théorique et conceptuel adapté à l'analyse critique et scientifique des processus, des significations[1] et mouvements des industries culturelles (et créatives), des industries de l'information ou des groupes de télécommunications (et des plateformes et de la *data*), des médias (traditionnels, nouveaux et numériques), des médiations car les Sic permettent l'analyse et l'interprétation des processus, logiques et stratégies liant technique, société et information au travers des réseaux, technologies et services de la communication.

Dans la première période, jusqu'à la fin du siècle passé, de la jeune discipline, une partie de ses concepts et de ses outils provient donc des disciplines antérieurement établies : la linguistique, la sémiologie, la cybernétique suivie de l'informatique, l'histoire des techniques, le droit, mais aussi l'économie et bien sûr la sociologie entre autres pour ses approches méthodologiques qualitatives au bénéfice d'une approche spécifique de l'information et de la communication. Cependant, la montée en puissance des technologies de l'information et de la communication d'abord, suivie de celles du numérique ensuite, a donné l'opportunité à la discipline de développer ses propres concepts d'usages, de médiation, de médias sociaux, d'identité numérique, etc. et d'instituer son propre champ riche d'approches et de méthodologies différentes et complémentaires portées par l'interdisciplinarité revendiquée. En essence, la raison d'être initiale des Sic a été de saisir les phénomènes à l'œuvre autour des médias de masse, de l'information et des pratiques sociales à partir des concepts d'information et de communication qui impliquent des dispositifs techniques, des industries, des plans d'action publique et des individus, tout particulièrement dans les zones d'ombres inexploitées. S'il s'agit dans les années 70 de la télévision, du journal télévisé, ce sont ensuite les NTIC, les jeux vidéo, les groupes industriels de communication, la communication des organisations qui sont étudiés et ainsi de suite au fur et à mesure des transformations lourdes et profondes des

1. En 1974, le comité des fondateurs avait un temps pensé nommer la future discipline, Sciences de la signification.

processus communicationnels et informationnels avec Internet, les plateformes, les industries créatives, etc.

Autrement dit, les Sic ont toujours questionné les processus, les acteurs, les productions, les significations, les dispositifs techniques au cœur des relations technique – société – information et communication, prises à la fois ensemble et sur un temps long. Posture épistémologique nécessaire pour comprendre et analyser les arrangements, les agencements, les structurations des réseaux, des acteurs, des dispositifs qui engendrent le passage : du réel à la connaissance, du fait à la représentation, de l'imaginaire à la technique, de l'objet à l'usage et de la *data* à l'information.

De la même façon, les Sic ont examiné le déploiement du numérique, qu'il soit pris comme une lignée technique autonome et spécifique (Paquienséguy, 2012), comme porteur d'idéologies ou de mutations conséquentes (Millerand, Proulx, Rueff, 2010). À leur habitude, elles le saisissent depuis plus de 30 ans déjà, au travers de la convergence des acteurs et des autoroutes de l'information par exemple, à la fois avec leurs outils méthodologiques et conceptuels habituels mais aussi dans une réflexion épistémologique et critique. Le « numérique » est porteur de bien des transformations et celles qui pèsent sur les humanités n'en représentent qu'une part comme la variété des travaux en Sic en témoigne[2]. C'est d'ailleurs cette posture ouverte, ancrée à la fois dans le terrain et dans l'épistémologie et tenue sur le temps long qui la rend plus que légitime à questionner le numérique, ses études et ses humanités.

Le numérique, omniprésent

Le numérique n'est pas le sujet en soi de cet ouvrage et nous ne chercherons pas ici à le définir d'un point de vue scientifique. Il suffit ici de constater dans tous les aspects de la société la généralisation de la structuration ou de l'articulation des phénomènes informationnels, communicationnels ou organisationnels à partir de technologies ou de procès numériques ou numérisés et de data, comme c'est le cas dans certains secteurs favoris de la discipline (éducation, savoirs, santé, politique, divertissements, culture, territoires, etc.). Elle l'a d'abord considéré dans la continuité de ses travaux et problématiques autour d'un axe sciences, techniques et société qui a principalement

2. Comme en témoigne l'ouvrage *Dynamique des recherches en sciences de l'information et de la communication*, accessible ici http://cpdirsic.fr/wp-content/uploads/2018/09/dynamiques-des-recherches-sic-web-180919.pdf.

abordé les logiques sociales et organisationnelles, la transformation des pratiques informationnelles et communicationnelles, les changements économiques et les représentations structurés et véhiculés par la technique numérique et les technologies qui en découlent.

La transition s'est manifestée par de multiples occurrences dont la plupart ont été étudiées par les Sic : le passage de l'analogique au numérique ; l'essor des bases de données et la transformation radicale de leurs méthodes d'indexation et d'interrogation ; le tuilage progressif et permanent qui a conduit des NTIC aux TICN (Paquienséguy, 2007) puis aux réseaux et médias sociaux (Stenger et Coutant, 2011) jusqu'aux objets connectés et aux data. Dans le même élan, nous en étudions les acteurs (Gafam, éditeurs, médias, innovateurs), les « effets », les usages, les pratiques, les mots, le sens, les représentations… (CPDirSic, 2018[3]). Et il en va de même pour les Humanités Numériques et les *Digital Studies* car encore une fois, les approches varient pour mieux saisir la complexité et l'intrication des processus analysés et théorisés comme le montre la deuxième partie.

Ainsi, comme pour tous les autres faits sociaux déjà étudiés et analysés en lien avec l'information et la communication, les Humanités Numériques constituent à la fois un cas d'études de terrain pour ce qui est des pratiques à l'œuvre, y compris celles de l'ensemble de la communauté scientifique, prises dans la continuité de ces pratiques ; et une suite à l'ensemble des travaux sur l'information et le document, depuis Paul Otlet (Otlet, 1989), Suzanne Briet (Briet, 1951), en passant par M. Buckland (Buckland, 1997) et Roger T. Pédauque (Pédauque, 2006).

L'informatique, mitoyenne

Le lien de l'information avec l'informatique et ce qui sera appelé plus tard l'Intelligence Artificielle (IA) se retrouve dès ses origines, souligné par les auteurs pionniers, Shannon et Weaver. Et même si l'information se distingue très tôt par sa polysémie et son caractère protéiforme, son développement trouve ses racines dans les approches mathématiques et computationnelles de « *The Mathematical Theory of Communication* » (Shannon et Weaver, 1949). Dès la publication de ce texte fondateur, l'analyse de l'information pose la triple problématique de la quantification, de la sémantique et de l'efficience qui a irrigué le courant de la recherche d'information ou *Information Retrieval*

3. *Idem.*

(Saracevic, 1999). Et on ne s'étonnera donc pas de la désignation de cette discipline tantôt au singulier : « la science de l'information », tantôt au pluriel : « les sciences de l'information ».

Malgré leurs différences bien établies, les approches françaises et anglo-saxonnes des sciences de l'information ont ceci de commun d'avoir développé leur épistémologie et leurs modèles théoriques sans jamais renier les apports de leur origine cybernéticienne (Ibekwé, 2012). L'ouvrage récent de Ronan Le Roux donne l'occasion de revenir sur cette filiation tout en donnant les clés de compréhension des raisons, en France, d'émancipation des Sic d'une part et du déclin de la cybernétique d'autre part au profit de l'automatique (Le Roux, 2018).

En matière d'émancipation, les sciences de l'information vont trouver dans le terrain des bibliothèques et de leurs pratiques professionnelles matière à formaliser des lois de puissance reposant sur le paradigme informationnel. Outre-Manche et Outre-Atlantique, les contributions de l'*Information Science and Librarianship* ont été réactivées dans les travaux des Sic françaises autour de la documentarisation du Web tenant compte d'un pendant médiatique négligé dans les travaux anglo-saxons (Pédauque, 2007).

L'augmentation exponentielle de la production de l'information survenue à partir des années 1960 (Price, 1962) et les problématiques corollaires de gestion et de traitement de quantités informationnelles toujours plus importantes constituent un enjeu majeur du développement d'une science qui cultivera un lien « nécessaire » d'abord avec la mécanisation, développée dans les pratiques professionnelles du monde des bibliothèques et à la bibliothéconomie pour gérer des objets documentaires *via* leur description bibliographique (métadonnées) ; puis avec l'automatisation et l'informatisation développées notamment grâce aux pratiques de la documentation (à partir de la fin des années 1960) qui, elle, favorisera le contenu des documents et leur description bibliographique et intellectuelle (métadonnées et indexation).

L'informatisation a apporté des solutions inédites pour les capacités de traitement et de stockage de l'information qui opèrent *via* des serveurs informatiques. Ce sont aussi de nouvelles méthodes de recherche d'information, automatisées, fondées sur des bases de données qui nécessitent des langages d'interrogation spécifiques conduisant à déconstruire l'information. Elles participent à introduire une redéfinition de l'information en l'équation « information = donnée + sens ». Cette nouvelle vision de l'information bouscule la notion

de document autant qu'elle introduit celle de données dans l'univers théorique de la discipline. Information et données se côtoient grâce à un seuil de réification, opérationnalisé par les Tic et les discours de la société de l'information (Labelle, 2005).

Le nouveau champ de recherche de l'information a connu un important développement, fécondé par des travaux issus de l'informatique, de la linguistique, du traitement automatique des langues (Bawden, 2012). En retour, les sciences de l'information ont nourri ces disciplines et leur ont permis d'intégrer des propriétés et/ou des méthodes propres. Pour illustrer ce dialogue et cette co-fécondation, nous retiendrons une question particulière, traitée par des recherches en Sic, et qui permet de souligner la dimension computationnelle des sciences de l'information et leur articulation sociale nécessaire aux sciences de la communication : les data et leur partage (data sharing).

La datafication comme phénomène sociétal, mais aussi comme processus informationnel et computationnel, a conduit à un développement des expertises autour des techniques de traitement, gestion et partage des données. Cette dernière fonction qui rassemble des enjeux et des acteurs de différentes natures, permise par des infrastructures (sectorielles, nationales ou européennes), devient la pierre angulaire d'une nouvelle économie – que certains courants qualifieront de nouvelle forme de capitalisme (Mirowski, 2008) – dont les mécanismes reposent sur le renouvellement des outils et des expertises réservées à leur exploitation dans des secteurs aussi différents que la finance, l'information scientifique, l'entreprise ou bien encore le débat politique.

Les sciences de l'information se penchent sur les data à travers le prisme des architectures et infrastructures numériques et/ou des standards ouverts. Les questions d'éditorialisation et de classification (voir ci-dessous) seront avancées pour prendre en charge les problématiques sémantiques de fouille de données, et de l'opacité des algorithmes et de leurs partis-pris économiques, voire politiques.

Science sociale, la discipline porte un regard attentif aux littéracies numériques cultivées au sein de cet environnement ; elles sont en vue de définir des curricula spécialisés. En effet, à partir du numérique, la discipline aborde et explore les liens qui existent entre technique et culture, en tenant compte des contextes d'appropriation et d'apprentissage qui les véhiculent.

Ce sont donc à la fois de nouvelles questions de recherches, mais aussi d'anciennes qui se réactualisent comme la recherche d'information

qui se mue en recherche de données et qui développe un second souffle aujourd'hui à l'aune de titres de revues lancées et de conférences internationales organisées. Les recherches entreprises se spécialisent plus que jamais et l'expertise nécessaire pour les appréhender, voire les évaluer, doit être à la mesure de la complexité des travaux entrepris par les chercheurs.

Une panoplie conceptuelle

Au-delà des études de terrain, qu'elles soient liées à des pratiques, à des technologies ou à des stratégies, au-delà de l'analyse des « effets » du numérique et au-delà de l'évolution des problématiques et méthodologies qui en sont affectées, l'apport des Sic à ce nouveau champ disciplinaire que seraient les *digital studies* se caractérise par les concepts et notions qu'elles proposent pour le saisir. L'approche la plus consensuelle sera de partir de la définition qu'elles donnent d'elles-mêmes : « Aujourd'hui, chaque discipline académique est affectée par les technologies numériques, mais nous ne possédons pas suffisamment de modèles théoriques pour étudier la façon dont elles influencent le savoir au niveau épistémologique (dans le sens de Gaston Bachelard). Nous pensons que cette question devrait être analysée à travers toutes les disciplines académiques, chacune d'entre elles devant maintenir sa différence tout en reconnaissant des affectations communes et des conséquences de cela au niveau épistémique (dans le sens de Michel Foucault). Cette idée est à la base de ce que nous nommons les "études digitales" ».[4]

Il y a 20 ans déjà, d'une part Yves Jeanneret se demandait « s'il y a[vait] vraiment des technologies de l'information ? » et Josiane Jouët proposait un « retour critique sur la sociologie des usages » montrant bien là qu'un tournant épistémologique de la discipline était à négocier en termes de concepts et de fondamentaux. Questionnant méthodiquement la montée en puissance des technologies de l'information, largement devenues technologies numériques, Jeanneret analysait ce que ces supposées nouvelles technologies apportent à la circulation sociale des informations et des savoirs. Pour lui, le mot « information », dont le sens paraît évident, masque toute la complexité des relations entre technologie et culture, et pour le cerner il mène un examen précis des propriétés réelles des « médias informatisés » qui soulèvent la question théorique et méthodologique de Jouët sur les formes de la production sociale dans des articulations

4. http://digital-studies.org/wp/fr

qui n'ont fait que se transformer et se renforcer entre « l'individuel et le collectif, le personnel et le social, le privé et le public » (Jouët, 2000 : 515). La question devient plus délicate et centrale encore si l'information se définit comme l'exploitation contextualisée ou l'éditorialisation des *data*.

Plusieurs concepts dégagent ces lignes de tension au fil du temps et ne pouvant ici nous attarder sur tous, nous en valoriserons quelques-uns parmi ceux qui témoignent des liens forts entre nos concepts historiques, pour certains identitaires, et les termes phares du contexte contemporain, qui relèvent des *digital studies* dont les Humanités Numériques semblent constituer un pan plus méthodologique qu'épistémologique autour des termes phares que sont : data, informatisation des procès, information ou éditorialisation. Notre discipline a effectivement franchi des seuils, a théorisé des ruptures, des changements de paradigme, de lignée, de modèles économiques, de logiques sociales… Et c'est ce que nous devons sans doute continuer à faire en nous posant la question suivante : comment les concepts et les notions que nous manipulons résistent-ils aux mutations du numérique ? Et ils résistent, car le numérique est une forme et non une substance.

Prenons donc quelques exemples éclairants qui nous sont déjà très familiers.

Médias

Les médias ont tout connu : ils ont été nouveaux, puis alternatifs, ils sont devenus sociaux. Les adjectifs varient sans refléter pour autant les changements profonds de ce qu'est un média ou une entreprise médiatique car la notion recouvre des réalités bien différentes depuis la naissance des *Media Studies* vers 1930. Par exemple, comment positionner les data compagnies comme *Alibaba* ou *Google* dans le registre médiatique ?

Usages

La question est centrale mais elle s'est transformée. Le paradoxe, ici seulement esquissé, est de constater aujourd'hui des usages massifs, à des échelles jamais imaginées, des mass-usages en quelque sorte, finalement assez convergents, d'objets connectés polyvalents très différents des mass-médias des années 70. S'ils vérifient sans conteste aujourd'hui la définition des usages sociaux proposée par Lacroix, Mœglin et Tremblay en 1992, ils s'accrochent plus à des applications ou à des services qu'à des terminaux, dans des mobilités et des

dispositifs qui valorisent les pratiques. Si le démantèlement du service public réclame sans doute de nouvelles problématiques adaptées à un contexte qui remplace l'usager par le consommateur, le primat de la connexion sur l'objet connecté est à l'œuvre dans la transformation des logiques sociales. Par exemple, comment les nouveaux médias sont-ils utilisés ? Comment les usages transforment-ils les stratégies et les productions industrielles, l'écriture médiatique, l'imaginaire et ses représentations ?

Conditionnalités du sens[5]

Pourtant centrale, l'étude des traces ne suffit pas à examiner les conditionnalités du sens car les points d'entrée se multiplient et se diversifient au fur et à mesure de la transformation des dispositifs d'écriture/lecture et du règne de la « trinité impériale : Performation/Prédiction/Préemption » (Noyer, 2013 : 2). Très souvent conjugués, plusieurs éléments (visibilités, créativités, matérialités, rationalités, sociabilités, textualités, autorités) sont nécessaires pour saisir et comprendre la conditionnalité du sens, par exemple, pour les problématiques de l'identité numérique, de l'*opendata*, ou de la narrativité y compris dans ses formes algorithmiques qui demanderaient de concevoir une « plurisémiotique » (Tétu, 2018).

Éditorialisation

Notion récente, arrivée à l'aube du numérique, l'éditorialisation a d'abord véhiculé une signification plus forte de la part des technologies numériques dans les modalités de production et diffusion des savoirs sur le Web. Mais surtout, l'éditorialisation, dans son développement et sa « propagation », en tant que démarche à part entière, rend compte aujourd'hui des recours aux outils et capacités du numérique pour élaborer et façonner des représentations ou des « visions du monde » (Vitali-Rosati, 2016). L'éditorialisation, comme concept et comme démarche, renvoie donc à un processus qui réussit à faire dialoguer potentialités du numérique, architectures conceptuelles et praxis documentaire, dans un but de co-construction. Ce dernier prend en compte la dimension sociale des pratiques qui interviennent pour modifier la production de la représentation en continu. C'est ce point en particulier qui en fait un concept important pour les Sic et leur filiation aux Humanités Numériques.

5. C'est avec émotion que nous nous appuyons ici sur les propos de Louise Merzeau lors de la première journée HN/DS de la SFSIC en mars 2016.

Information, document, connaissance

Cet ensemble regroupe trois concepts historiques qui ont contribué à fonder le versant « informationnel » de la discipline. Les travaux menés autour de leur définition et de leurs théories ont animé – et animent toujours – un pan important des recherches en sciences de l'information. Ces travaux présentent des pendants empiriques dans le milieu de l'entreprise, à travers des problématiques de gestion de l'information, du document, des connaissances, souvent plus connus par leurs formules anglicistes : *Information Management, Document Management, Knowledge Management* (KM).

Dans le domaine de l'information scientifique, on observe le développement de techniques numériques de fouille, d'extraction et d'analyse textuelle (*text mining*). Elles ont permis de développer de nouveaux fronts de recherche qui ont ouvert des perspectives inédites en matière de traitement de l'information. Au-delà des métadonnées traditionnellement étudiées, les travaux menés autour de la structure du document et de sa dimension textuelle redéfinissent l'exploitation des documents en offrant de nouvelles capacités d'extraction d'une information toujours plus volumineuse et dont la production est amenée à s'accélérer (Bertin, 2016).

Ces techniques permettent de penser une nouvelle synergie des concepts d'information, de document et de connaissance. Adoubés puis décriés les uns après les autres, selon la période où les sciences de l'information se sont développées puis imposées comme discipline, ces trois concepts cessent d'être pensés en dualité. À travers les techniques numériques du *Data Mining* et du *Deep Learning* (affiliées à l'Intelligence Artificielle), les trois concepts sont désormais traités par le biais de la « relation », du « continuum ». Le champ de l'information scientifique et technique représente un terrain particulièrement approprié pour cette nouvelle vague de travaux qui incarne l'apport des sciences de l'information aux méthodes numériques.

Classement, classification

Rattachés à la genèse des sciences de l'information, le classement et les classifications sont convoqués pour affronter un nouveau cycle de massification de la production de l'information et de la donnée. Ces concepts sont au cœur des enjeux des pratiques sociales académiques et/ou citoyennes qui prennent part à la sélection, à la description et à la classification des contenus (photos, tweets…). Leurs caractéristiques intellectuelles sont questionnées à l'aune de dimensions « émotives »,

« créatives »… qui donnent lieu à de nouveaux terrains d'observation. Les perspectives d'analyses prennent en compte les plateformes dédiées, les acteurs impliqués et leurs stratégies, mais aussi les tensions (sociales, culturelles, politiques) en jeu dans de nouvelles formes et une plus grande diversité de « mise en ordre » du monde (Liquète, 2013).

Chaîne de la valeur

Le concept de la chaîne de la valeur, formalisé dans les années 1970, et alors que les services d'information connaissaient un grand développement, se conjugue désormais au pluriel pour prendre en compte les nouveaux modes de création de valeur qui font appel à la *data* ou au *Big data*. Ces modes passent le plus souvent par des dispositifs de médiations dynamiques de production de services « innovants » proposés le plus souvent sur des plateformes collaboratives et participatives. Les recherches attentives à la création de la valeur dans les économies contemporaines s'orientent vers l'analyse et la compréhension des mécanismes et des facteurs associés à l'exploitation de la donnée pour le développement de services à valeur ajoutée. Ces recherches fondées le plus souvent sur des travaux empiriques tentent de développer en parallèle une réflexion plus théorique en lien avec les problématiques de risque et d'incertitude, propres à la production de la valeur des services d'information.

Cette panoplie conceptuelle n'a bien entendu qu'une vocation illustrative. Elle a le mérite de donner un aperçu de la manière avec laquelle le numérique interroge les concepts, l'étendue de leur résistance ou de leur résilience. Leur capacité d'accueil du numérique dessine en filigrane les enjeux du projet scientifique de la discipline, relevant moins d'un objet défini, mais plutôt d'un projet anthologique (Jeanneret, 2004).

Conclusion

Le cœur de compétences des Sic revient sans doute à observer et interpréter la réalité sociale (ou les formes sociales, Jouët) dans ses productions et ses représentations à travers ses acteurs, historiques et *pureplayers*, industriels, collectifs et individuels, au fil de leurs logiques, stratégies et pratiques. À elles de concevoir les outils de l'interprétation et donc « de penser le numérique » (Merzeau, 2016). Depuis ses origines, notre discipline articule l'information et la communication de façon indéfectible car c'est cette articulation qui fait sens au gré des discours, des transformations, des alliances, des

évolutions techniques. Aujourd'hui, les deux termes connaissent eux-mêmes un changement profond en raison des technologies du numérique, mais surtout des acteurs qui les exploitent à l'échelle industrielle et qui en usent à l'échelle individuelle.

Le tournant digital valide, s'il en était besoin, l'articulation entre information et communication et la rend sensible, véritablement, à la suite des discours antérieurs (politiques et critiques) sur la société de l'information, le capitalisme informationnel, l'emprise de la communication, le paradigme relationnel qui favorise la médiation ou l'économie politique de la communication qui l'inscrit totalement dans les industries de l'information et des contenus. Ces changements rendent caducs certains cloisonnements antérieurs qui ont été structurants, y compris pour les Sic. À titre d'exemple, la distinction ancienne production-émission-réception, la typologie des médias, les socio-types, les différents modèles économiques, les codes et signifiances. Dans le même temps, ils valorisent la chaîne qui fait de la *data* une information formatée pour la communication.

De même, la massification des données oblige les sciences de l'information à examiner le « nouvel ordre informationnel » (Miège, 2000) dans lequel elles baignent et la stratégie des acteurs qu'elle côtoie. De manière inédite, la discipline construit avec la communication des problématiques partagées lui permettant de mieux comprendre la nature de son propre environnement (une industrie du contenu) et les techniques numériques qu'elle mobilise en tant qu'objets sociaux (ex. algorithmes). Ainsi, les sciences de l'information ne renoncent pas à leurs racines computationnelles et mathématiques, mais acquièrent une meilleure compréhension des environnements sociaux dans lesquels elles s'insèrent. Elles deviennent ainsi les mieux placées pour analyser les données en vue de produire du sens.

Retravailler les termes fondateurs, information et communication, penser les concepts et méthodes adéquats à l'analyse critique d'une réalité sociale, économique et politique différente, contribue à la fois à renforcer les Sciences de l'Information et de la Communication et leur légitimité à continuer à apporter des réponses à ces transformations liées au numérique, comme elle le fait depuis les années 1980. Nous soutenons ici l'idée d'une intégration naturelle des Sic dans les *Digital Studies* car depuis leurs origines, elles travaillent à interroger les développements du numérique, à le penser et à décrire cet environnement, pour le transformer en le pensant.

POSITIONNEMENT : 3 PROPOSITIONS

Ce n'est pas parce que le contexte change que la problématique doit changer car l'usage, l'écriture, l'imaginaire restent des constantes ; pour l'analyse des stratégies, des significations comme pour celles des productions.

Toutes les disciplines sont aujourd'hui concernées par le numérique. Il questionne leur épistémologie et l'ordre de leur savoir. Or, les Sic, leur histoire, leurs objets de recherche et leur interdisciplinarité, leur donnent assurément un avantage sur la compréhension des enjeux qui se jouent autour du territoire des Humanités Numériques. Nos trois propositions sont formalisées à partir de cet avantage :

– Proposition 1 : Un devoir de pratique

À l'instar d'autres disciplines, les Sic font appel aux méthodes numériques pour saisir, analyser, comprendre, dans quelle mesure elles transforment les pratiques intellectuelles. Pratiquer les méthodes des Humanités Numériques permet aux Sic de poursuivre leur projet d'analyse des liens entre la technique et la société, ou bien de la compréhension du poids des techniques sur la culture. L'un des apports majeurs des Sic dans cette démarche est qu'elles sont en mesure de prendre en compte les apports (conceptuels, théoriques) récents des sciences à l'origine du numérique ; elles en connaissent les origines, les spécificités et sont les mieux placées pour distinguer ce qui relève de l'innovation ou du renouvellement.

– Proposition 2 : Un devoir de critique

Les Sic ont un devoir de critique vis-à-vis des Humanités Numériques, à l'instar de celui qu'elles ont eu vis-à-vis d'autres débats (ex. société de l'information, pour ne citer que celui-là), d'autres territoires. Une critique épistémologique, sociale et empirique, fondée sur une connaissance et une maîtrise approfondie et réelle du numérique comme technique relève donc de la responsabilité des Sic. Cette critique sera d'autant plus fondée qu'elle prendra en compte les acteurs du numérique, les stratégies qu'ils mobilisent, pour mieux rendre compte de sa matérialité. Il s'agit donc de montrer les apports du numérique aux Humanités mais aussi leur dysfonctionnement, leurs risques.

– Proposition 3 : Un devoir de formation

Enseigner le numérique est une autre des responsabilités des Sic. L'enseignement et l'apprentissage des méthodes, des langages, des standards, des protocoles, des formats rattachés au numérique sont indispensables pour appréhender les Humanités Numériques, leur « éthique », leur « épistémologie », voire leur « morale ». Ce n'est qu'au prix de cet effort que les Sic pourront poursuivre une critique et un débat argumentés autour des Humanités Numériques.

BIBLIOGRAPHIE

Bawden D., Robinson L., 2012. *Introduction to Information Science*. Facet Publishing.

Bertin M., Atanassova I., Gingras Y., Larivière V., 2016. « The invariant distribution of references in scientific articles ». In *Journal of the Association for Information Science and Technology* (*JASIST*), 67 (1), pp. 164-177.

Briet S., 2019. *Qu'est-ce que la documentation ?* Éditions documentaires industrielles et techniques. Paris, 1951. Disponible à partir de : http://martinetl.free.fr/suzannebriet/questcequeladocumentation/briet.pdf. Consulté le 25 novembre 2019.

Buckland M., 1997. « What is a document? » In *Journal of the American Society for Information Science* (*JASIST*), *48*(9), pp. 804-809.

Coutant A., Stenger T. (dir.), 2011. « Ces réseaux numériques dits sociaux », *Hermès*, n° 59.

Ibekwe-Sanjuan F., 2012. « The French conception of information science: "une exception française"? » in *Journal of the American Society for Information Science and Technology* (*JASIST*), *63*(9), pp. 1693-1709

Jeanneret Y., 2000. *Y-a-t-il vraiment des technologies de l'information ?* Lille, Presses du Septentrion.

Jouët J., 2000. « Retour critique sur la sociologie des usages ». In *Réseaux*, vol. 18, n° 100. *Communiquer à l'*ère des réseaux. pp. 487-521.

Labelle S., 2001. « La "Société de l'information" à décrypter ». In *Communication & Langages*, (128), pp. 65-79.

Lacroix J.-G., Mœglin P. et Tremblay G., 1992. « Usages de la notion d'usage : NTIC et discours promotionnels au Québec et en France », *Les nouveaux espaces de l'information et de la communication*, in *Actes du 8e Congrès national des sciences de l'information et de la communication, Lille, mai 1992*.

Le Roux R., 2018. *Une histoire de la cybernétique en France* (*1948-1975*), Paris, Éditions Classiques Garnier.

Liquète V., Kovacs S. (dir.), 2013. « Penser, classer, contrôler », revue *Hermès*, n° 66, CNRS Éditions.

Miège B., 1989. *La Société conquise par la communication*, Grenoble, PUG.

Miège B., 2000. *Les Industries du contenu face à l'ordre Informationnel*. Grenoble, PUG.

Mirowski Ph., 2008. « The commercialization of science and the response of STS ». In E. Hackett, O. Amsterdamaska, M. Lynch &

J. Wacjman (eds.), *The Handbook of Science and Technology Studies* (pp. 635-689). MIT Press.

Noyer J.-M., 2013. « La transformation numérique : quelques procès en cours », *Revue française des sciences de l'information et de la communication* [En ligne], 2 | 2013, mis en ligne le 1er janvier 2013. URL : http://journals.openedition.org/rfsic/377; DOI : 10.4000/rfsic.377. Consulté le 25 novembre 2019.

Otlet P., 1989. *Traité de documentation, le livre sur le livre : théorie et pratique*. C.L.P.C.F, Nouvelle édition.

Paquienséguy Fr., 2012. « L'usager et le consommateur à l'ère numérique » in Vidal G. (dir), *La sociologie des usages : transformations et continuités* Éditions Hermès Lavoisier, collection Traités, pp. 179-212.

Paquienséguy Fr., 2007. « Comment réfléchir à la formation des usages liés aux technologies de l'information et de la communication numériques ? », *Les Enjeux de l'information et de la communication*, 2007/1, p. 63-75.

Pédauque Roger P., 2006. *Le document à la lumière du numérique*. C & F éditions.

Pédauque Roger P., 2007. *La redocumentarisation du monde*. Paris, Cépaduès.

Price, D. J. de Solla, 1963. *Little Science, Big Science*. Columbia University Press.

Saracevic T., 1999. « Information science » in *Journal of the American Society for Information Science & Technology* (*JASIST*), n° 50 (12), pp. 1051-1063.

Shannon Cl., Weaver W., 1949. *The Mathematical Theory of Communication, University of Illinois.*

Tétu J.-Fr., 2018. *Le récit médiatique et le temps accélérations, formes, ruptures*, Paris, L'Harmattan.

Vitali-Rosati M., 2016. « Qu'est-ce que l'éditorialisation ? ». In *Sens public*, 2016. Disponible à partir de https://hal.archives-ouvertes.fr/hal-01599208. Consulté le 25 novembre 2019.

PARTIE II

Humanités Numériques : de l'éthique au politique

INTRODUCTION PARTIE II

Humanités Numériques
De l'éthique au politique

La place et les enjeux liés aux Humanités Numériques invitent à renforcer une réflexion critique afin de situer et contextualiser à la fois les discours et les méthodes et ainsi contribuer à la production de connaissances. Cette approche critique permet d'appréhender les multiples conditions d'émergence d'« une société numérique », objet de nombreux questionnements dont s'emparent les chercheuses et chercheurs en sciences de l'information et de la communication. De ce point de vue, les Sic qui examinent à la fois les conditions de production du sens, tout autant que les dispositifs et les formes de médiation contribuent à la réflexion sur les enjeux d'une herméneutique numérique, sur les conditions politiques, économiques et éthiques de la production scientifique. Aussi, cette deuxième partie privilégie quatre principaux axes de réflexion autour de ces enjeux : celui des discours, du positionnement éthique, des liens entre computation et interprétation et enfin des conditions de possibilité du sens au regard de l'économie de la production scientifique.

L'examen des discours politiques et publics sur les Humanités Numériques, en particulier ceux qui portent sur leur enseignement, atteste des liens entre humanités et informatique et pose la question des conceptions de la régulation que ces discours induisent. Situer ces discours constitue une des conditions de la production scientifique et rend compte des résonnances avec de précédentes visions systémiques, relatives par exemple aux discours sur les « autoroutes de l'information », voire antérieurement, aux discours sur la cybernétique. Fortes des expertises sur ces discours, les Sic

sont à même de développer de nouveaux concepts, tels que celui de « trace » qui propose un regard neuf et inédit sur l'environnement digital (**chapitre IV**).

Cet environnement digital invite en effet les chercheurs en Sic à penser un positionnement éthique et nécessite une attention renouvelée à l'intégrité scientifique, cela à chaque étape de la recherche, de la production jusqu'à la restitution des résultats. Ce questionnement et cette réflexivité sur sa propre pratique scientifique ne sont pas nouveaux dans la discipline. Ils se trouvent néanmoins réactualisés au regard des Humanités Numériques et surtout des nouveaux enjeux qui les accompagnent avec la manipulation d'outils et de méthodes numériques (ou numérisées) et l'usage de volumes importants de données. Dès lors, les Sic françaises par leur interdisciplinarité, leur ouverture à l'international et la permanence de leur questionnement offrent un espace de réflexion privilégié pour accueillir la question de l'éthique au sein de leurs démarches collectives. En témoigne la création récente du groupe GENIC au sein de la SFSIC (**chapitre V**).

La recherche du sens dans le rapport aux objets techniques est également majeure et réinvestie au prisme des Humanités Numériques, avec de nouvelles notions qui peuvent là encore être mobilisées dans les Sic. Sur cet aspect, la rencontre entre les Sic et les études digitales amène à explorer de nouveaux concepts, comme celui de « computation-interprétation » ou d'« hyperdocument » dans toutes ses déclinaisons (hyperdocumentalité, hyperdocumentarité). Ceci pour s'autoriser à penser le format digital dans l'articulation entre numératie et littératie ou encore entre information et communication (**chapitre VII**).

Ainsi les Sic, par leur multidimensionnalité, leur interdisciplinarité et leur inventivité, sont à même d'apporter quelques éléments de réponse aux nouveaux défis posés par les « Humanités Numériques ». C'est ce que nous invite à explorer toute cette deuxième partie.

Christine Barats
Université de Paris, Cerlis, UMR 8070

Ghislaine Chabert
Université Savoie Mont Blanc, LLSETI EA 3706

CHAPITRE IV

Situer et contextualiser les discours publics et politiques

Christine Chevret-Castellani
Université Sorbonne Paris Nord, LabSic

Résumé : Le syntagme « Humanités Numériques » constitue désormais un « label » promu par des acteurs publics et politiques, en particulier dans le domaine de l'enseignement et de la recherche. Il renvoie à un contexte politique, sinon idéologique, à interroger. Pour ce faire, il s'agit de situer ces types d'« humanités » non seulement dans ce contexte politique mais aussi de les envisager comme contribuant à l'édification de ce contexte déterminée par une approche systémique de l'environnement numérique. Certains discours d'acteurs publics et politiques français tendent en effet aujourd'hui à la promotion d'un enseignement de l'informatique comme lié à une manière, non seulement de construire des enseignements, d'instruire et de former, mais aussi de projeter une « société numérique » dans laquelle se déploieraient des interactions entre l'homme et la machine.

Mots-clés : contexte politique, discours publics et politiques, régulation, cybernétique, trace.

Abstract: The syntagm «digital humanities» is now a «label» promoted by public and political actors in the field of education and research. It refers to an ideological context as to be questioned. In fact, it is a question of situating these «humanities» not only in this context but also to consider them as contributing to the construction of this context determined by a systemic approach to the digital environment. Some of French public and political actors speeches today tend indeed to promote a conception of «digital society» in which interactions between human and machine should deployed.

Keywords: Political context, Public and political speeches, Regulation, Cybernetics, Trace.

« [...] Notre futur est celui des années 40 »
Philippe Breton, *L'Utopie de la communication*, 2009, p. 10.

L'expression « Humanités Numériques » apparaît dans une publication anglaise de 2008 (Mounier, 2018). Elle renvoie à l'idée de mettre l'informatique au service du traitement de corpus textuels. Pierre Mounier explique en effet que les « Humanités Numériques » seraient nées après la Seconde Guerre mondiale d'un partenariat entre le Père Busa, qui cherchait à faire un index de la *Somme théologique* de Thomas d'Aquin, avec l'entreprise informatique IBM, inaugurant par là même une association des humanités classiques avec le monde industriel (Mounier, 2018). Entre 2010 et 2016, cette expression se démocratise ; des médias comme *Le Monde*[1] ou *The Guardian*[2] s'en emparent pour désigner des objets d'études liés au numérique (Valluy, 2015). Elle désigne aujourd'hui non seulement un domaine de recherches et de débats sur la place du numérique, mais suppose aussi une injonction à des changements de méthodes de recherches liés au développement des technologies de l'information et de la communication.

Comme l'indique le texte d'orientation du réseau « Humanités Numériques plurielles », le syntagme « Humanités Numériques » constitue désormais un label promu par les pouvoirs publics dans le domaine de l'enseignement et la recherche. Il induit des « prescriptions en faveur d'actions publiques favorisant l'hybridation

1. « Qui a peur de l'Open access ? », *Le Monde*, 15 mars 2013.
2. Preston Alex, « The war against humanities at Britain's universities », *The Guardian*, 29 mars 2015.

de l'informatique et des humanités » (Valluy, 2018). Dès lors, les « Humanités Numériques » deviennent en elles-mêmes un contexte politique, sinon idéologique, à interroger. Comment ce contexte politique se construit-il ? Quelles sont les conceptions des « Humanités Numériques » véhiculées dans les discours des pouvoirs publics et des acteurs politiques ? Comment évoluent-elles ?

L'idée qui gouverne ce chapitre est de situer les « Humanités Numériques » non seulement dans un contexte politique mais aussi de les envisager comme contribuant à la construction de ce contexte. C'est la raison pour laquelle nous examinons une conception politique dominante (Bouquillion et Matthews, 2010) ancrée dans une approche systémique de l'environnement numérique dans lequel se déploieraient des interactions entre l'homme et la machine. En effet, les discours des acteurs publics français ont tendance à envisager l'enseignement informatique d'aujourd'hui comme lié à une manière non seulement de construire des enseignements, d'instruire et de former, mais aussi de projeter une « société numérique », c'est ce que nous envisagerons dans une première partie.

Un certain type de citoyenneté qui impliquerait la capacité à intervenir sur Internet[3] et à en modifier, sinon la gouvernance du moins la configuration, suppose une formation au métier de développeur de telle façon que ces spécialistes de l'informatique participent à la fois au déploiement de la « French Tech » et à l'invention de systèmes intégrant la régulation. Cette dernière peut aujourd'hui être définie comme un dispositif techno-juridique (Dulong de Rosnay, 2016, p. 14). Le code informatique est en mesure de faire intégrer des règles automatiquement par l'architecture des plate-formes tandis que le droit propose des règles juridiques (conventions internationales, directives et règlements européens, loi et décrets, etc.). D'un côté, des acteurs publics français, comme ceux des autorités administratives indépendantes, vont considérer la réglementation comme nécessaire pour réguler, tout en cherchant également de plus en plus d'alternatives. D'un autre côté, des acteurs privés du numérique envisagent la régulation par le code informatique comme pouvant se substituer à la loi, celle-ci étant vécue comme un frein à l'innovation. C'est ce que nous verrons dans une deuxième partie.

En déployant ses potentialités d'intervention, consubstantielles de compétences spécifiques, le citoyen – et ce qui ressort de certains

3. Par l'orthographe « Internet », nous désignons l'outil et par « l'internet » nous signifions le média.

discours politiques que nous évoquerons – serait en mesure de devenir régulateur de lui-même. Cette conception de la régulation, telle qu'envisagée aujourd'hui, trouve son point d'ancrage dans la cybernétique de Norbert Wiener dont les fondements avaient déjà ressurgi, en 1994, dans les discours sur « les autoroutes de l'information » du vice-président états-unien Al Gore. Enfin, elle apparaît comme le point d'aboutissement d'une évolution de la conception ayant présidé à la construction du droit de l'internet qui implique notamment de véhiculer un langage informatique – à l'instar du terme de « données » – et de s'orienter vers une patrimonialisation individuelle des contenus du web. C'est ce que nous aborderons dans une troisième partie.

Les « Humanités Numériques » dans les discours des acteurs publics et politiques français : bâtir une « société numérique »...

Entre 2013 et 2016, le Conseil National du Numérique (CNNum) publie deux rapports au sujet de l'enseignement des « Humanités Numériques », le premier concernant l'enseignement secondaire et le second portant sur les formations du Supérieur, dont les finalités sont de construire une culture propre à une « société numérique ».

Les fondements d'un enseignement en « Humanités Numériques » dans les rapports publics

Intitulé « Jules Ferry 3.0, Bâtir une école créative et juste dans un monde numérique », publié en 2014 et faisant état de 40 recommandations formulées en octobre 2013, le premier rapport vise à promouvoir un baccalauréat spécifique à cette discipline. Il s'agit tout d'abord d'intégrer l'enseignement dans une certaine modernité : « Ce nouveau bac s'inscrirait dans son époque. Il refléterait l'aventure de la jeunesse et revitaliserait les études secondaires avec la création numérique, le design mais aussi la découverte des big data, de la datavisualisation, des métiers informatiques et créatifs. »[4]. Cette spécialité du baccalauréat aurait pour vocation de générer une nouvelle motivation chez les élèves et les enseignants :

> Ce nouveau bac est par essence le plus général des bacs généraux puisqu'il se positionne au croisement des sciences, lettres, et sciences humaines et sociales, en décloisonnant ces champs du

4. Conseil National du Numérique, « Jules Ferry 3.0, Bâtir une école créative et juste dans un monde numérique », 2014, p. 9.

> savoir. Son but est de revisiter les humanités dans toute leur richesse et leur modernité, en s'appuyant sur les sciences, les techniques et les pratiques du numérique. Il est tout à la fois symbole, catalyseur, banc d'essai : l'informatique et le numérique sont essentiels pour tous les enseignements. Montons au cœur des lycées un nouveau bac qui rassemblera les appétits des élèves et donnera un souffle de liberté à de nombreux professeurs. (CNNum, 2014, p. 17).

Les recommandations du CNNum insistent sur cette nouvelle motivation liée à un enseignement novateur : « Il ne s'agit rien de moins que d'inventer "le bac de l'individu créatif de la civilisation numérique". »[5]. Les objectifs restent cependant de préparer aux exigences économiques[6], telles que les préconisent les politiques européennes, notamment depuis le Livre Blanc de Jacques Delors approuvé par le Conseil européen le 11 décembre 1993[7], sur le développement des enseignements en technologies. Les contenus de ces enseignements « novateurs » sont envisagés à la fois comme devant permettre de développer des compétences scientifiques et techniques, mais aussi de construire une culture numérique plus générale, comme l'illustre le rapport publié sur l'intelligence artificielle par Cédric Villani en 2018[8]. Dans le cadre de la nouvelle discipline « Humanités Numériques », les cours du Lycée doivent à la fois renforcer les compétences en algorithmique et faire émerger puis consolider un type de socle commun autour d'une culture spécifique dans une « société numérique »[9]. À partir de 2019, dans le baccalauréat qui fait l'objet d'une réforme jusqu'en 2021, réforme intitulée « horizons 2021 », le terme « humanités » a été introduit dans un enseignement de spécialité « humanités, littérature et philosophie » et a été séparé de l'enseignement du numérique devenu un autre enseignement de spécialité « numérique et sciences informatiques »[10]. Dans le cadre de cette réforme, l'association des champs disciplinaires « humanités » et « numériques » a donc finalement été abandonnée.

---

5. *Ibid*., p. 50.
6. *Ibid*., p. 53.
7. Livre blanc de la Commission, « Croissance, compétitivité, emploi », http://www.cvce.eu/obj/livre_blanc_de_la_commission_croissance_competitivite_emploi_1993-fr-b0633a76-4cd7-497f-9da1-4db3dbbb56e8.html (consulté le 2 mai 2019).
8. Villani Cédric, *Donner un sens à l'intelligence artificielle. Pour une stratégie nationale et européenne*, La Documentation française, 2018.
9. *Ibid*., p. 121.
10. Horizons 2021, http://www.horizons2021.fr/ (consulté le 2 mai 2019).

Depuis 2016, dans l'enseignement supérieur et la recherche, ces champs disciplinaires font l'objet de préconisations eu égard à l'évolution des diplômes délivrés, en particulier par les Universités. Ainsi, en mai 2016, le CNNUm a remis ses travaux sur le sujet à Thierry Mandon, alors Secrétaire d'État, dans un second rapport succédant à celui de 2014 et inscrivant une continuité entre les enseignements secondaire et supérieur. Ses recommandations reposent sur 12 points. Il s'agit notamment d'« ouvrir les accès à la littératie numérique et proposer des formations diverses, des sciences informatiques aux cultures numériques, aux Humanités Numériques et au "pouvoir d'agir", pour tous les acteurs de l'Enseignement supérieur et de la recherche, étudiants, enseignants-chercheurs, administratifs. »[11]. Les recommandations du CNNum soulignent également l'apport de ces formations :

> [...] Les sciences de l'homme et de la société, en questionnant les usages du numérique et les transformations sociales qui y sont liées, permettent de penser la manière dont les personnes ne se contentent pas d'être des utilisateurs des techniques qui leur sont proposées, et d'observer comment ils se les approprient, les détournent, inscrivent ces pratiques dites révolutionnaires dans la grande histoire des hommes [...] (CNNum, 2016, p. 10).

Revenons alors sur deux traits saillants de ces discours publics. Le premier trait porte sur la justification d'un enseignement spécifique, par lequel une « reconnexion »[12] de l'enseignement avec la jeunesse est espérée. Pierre Mounier indique que par l'introduction des « Humanités Numériques » dans les formations « [...] il s'agit en l'occurrence de remédier, par une idée simple, à la désaffection et la dévalorisation qui frappent les filières de formation littéraire depuis plusieurs décennies » (Mounier, 2018, p. 11). On peut également concevoir que ces formations étant dévalorisées économiquement et socialement, sont désaffectées par les étudiants. Par ailleurs, l'adaptation de ces formations aux exigences économiques n'est pas un élément anecdotique. Rappelons sur ce point qu'il y a dans « Europe 2020 » – la stratégie de croissance de l'Union européenne – une priorité relative au développement d'une « économie fondée sur la connaissance et l'innovation », présentée sur le site du ministère

11. Conseil National du Numérique, « Une nouvelle dynamique pour la métamorphose numérique de l'Université », Dossier de presse, 2016, p. 5.
12. Conseil National du Numérique, « Jules Ferry 3.0, Bâtir une école créative et juste dans un monde numérique », 2014, p. 53.

de l'Enseignement supérieur, de la Recherche et de l'Innovation[13]. Le deuxième trait concerne la promotion d'une nouvelle « société numérique ». Cette dernière – qui justifie un diplôme en « Humanités Numériques » – renvoie à l'usage d'une expression qui n'est pas sans rappeler la manière dont la « société de l'information » (Labelle, 2001) avait précédemment été promue.

Les fondements de la « société numérique » dans les discours politiques sur la « société de l'information »

Cette dite « société de l'information » a fait l'objet, en France, d'un programme d'action gouvernemental (PAGSI), promu dès août 1997 par Lionel Jospin à l'Université d'été de la communication d'Hourtin, en Gironde[14]. En faisant alors de l'entrée de la France dans la « société de l'information » l'une des priorités de l'action gouvernementale, le premier ministre de l'époque a fixé des objectifs en termes de transformation de la société. Comme dans les rapports publics du CNNum de 2014 et 2016, le PAGSI se donnait notamment comme objectif de changer l'école. En 1998, il s'agissait d'effectuer cette transformation en y intégrant les « technologies de l'information », par la mise à disposition de matériels et la facilitation de l'accès à l'information. Il est alors aussi question d'améliorer l'administration, de favoriser la « nouvelle économie », le développement de l'usage de l'internet au quotidien et l'adaptation au droit.

C'est cette adaptation du droit qui fera l'objet d'une deuxième étape liée au PAGSI, étape annoncée à Hourtin le 26 août 1999, et qui vise essentiellement la lutte contre les nouvelles formes de criminalité, l'adaptation juridique concernant la liberté de communication, la sécurisation des contenus et le droit d'auteur[15]. Cependant, lors de cette Université d'été, les discours se révèlent polyphoniques. Ainsi, Catherine Trautmann, alors ministre de la Culture et de la communication, considère que personne ne détient le monopole des normes de régulation de l'internet à l'échelle mondiale, mais surtout que la régulation de l'audiovisuel est plus importante. L'internet étant

13. Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation, http://www.enseignementsup-recherche.gouv.fr/cid71587/la-strategie-europe-2020-pour-une-croissance-intelligente-durable-et-inclusive.html (consulté le 2 mai 2019).

14. Jospin Lionel, « Discours prononcé lors de l'inauguration de l'Université de la Communication : préparer l'entrée de la France dans la société de l'information », Archives du premier ministre Lionel Jospin (25 août 1997), http://archives.gouvernement.fr/jospin_version3/fr/ie4/contenu/5519.htm (consulté le 14 février 2021).

15. Commission française pour l'Unesco, *La « société de l'information » : glossaire critique*, Paris, La Documentation française, 2005, p. 19.

limité à des possibilités de communiquer, de numériser du patrimoine ou au mieux de favoriser la créativité. La confrontation avec les difficultés à envisager une régulation mondiale l'amène à déclarer : « Je ne crois pas que l'internet soit à même de remplir l'ensemble des fonctions qu'assurent les médias dits traditionnels, fonctions non seulement techniques mais aussi politiques et sociales. »[16].

Ainsi, la deuxième étape liée au PAGSI laisse place à la « nouvelle société de l'information ». Cette dernière améliorerait la nature du travail en favorisant la créativité, l'intégration sociale et la participation politique. Sa promotion renvoie à une représentation du monde, liée à la nécessité de transformation de la société industrielle, mais aussi à une conception de la communication favorisant l'inclusion et le développement de formes de démocratie participative. L'expression « information » est cependant remplacée par « numérique » dès 2004, dans le cadre de la loi du 21 juin relative à la confiance en l'économie numérique. En 2004, la circulation de l'information *via* les réseaux est déjà une réalité ; désormais la société doit « se numériser ». La manière de concevoir l'adaptation du droit à l'internet a également évolué depuis. C'est ce que nous allons voir à travers la promotion par les pouvoirs publics d'alternatives au droit, à l'encadrement juridique et à la réglementation.

...Où le citoyen deviendra régulateur de lui-même

Pour envisager une régulation, semble s'imposer aujourd'hui, pour certains acteurs publics français mais surtout politiques, la nécessité de recourir à d'autres moyens que la loi. L'apprentissage de compétences en « Humanités Numériques » et en informatique doit en effet conduire à former des développeurs capables de réguler eux-mêmes des systèmes. Il représente un outil pour le développement économique, en particulier lié à la capacité d'innovation de la « French Tech », mais également un objectif idéologique, celui d'une France en mesure d'imposer en Europe, encore dominée par un modèle états-unien, ses propres conceptions – celles qui impliquent en particulier davantage de protection des données personnelles – de la régulation. Cette dernière est également promu de manière à répondre à un

16. Trautmann Catherine, « Déclaration sur le développement de la société de l'information dans le contexte de la mondialisation, notamment dans l'audiovisuel public, la mise en place des réformes dans les télévisions publiques et le projet de mondialiser l'espace de la communication publique », Hourtin, 23 août 1999, *Vie-publique*, Collection des discours, https://www.vie-publique.fr/discours/208562-declaration-de-mme-catherine-trautmann-ministre-de-la-culture-et-de-la (consulté le 14 février 2021).

objectif politique, celui annoncé au citoyen de parvenir à être sur Internet le régulateur de lui-même, comme il a pu l'être en adoptant des règles de netiquette, à l'instar des Requests for Comments (RFC) dans les années 1990, mais aussi en se responsabilisant davantage notamment en prenant conscience des risques (usurpation d'identité, vol de données, piratage de comptes, etc.) et en étant actif dans ses paramétrages lors de sa navigation sur le web.

Une autorégulation comme alternative à la loi

Ainsi, Jean Lessi, secrétaire général de la Commission nationale de l'informatique et des libertés (CNIL), insiste lors de la manifestation « Mise en œuvre du RGPD en France, quel État ? »[17], sur cette idée que la régulation n'est pas réductible au droit. L'idée que les pouvoirs publics peuvent trouver des alternatives à la loi pour réguler l'internet n'est pas récente. C'est à cette fin de médiations que le Forum des droits sur l'internet (FDI) avait été créé en 2001. Association financée à 90 % par l'État, le forum avait pour vocation de contribuer au dialogue entre acteurs publics et acteurs privés de manière à construire une co-régulation[18]. Ces alternatives à la loi ont cependant évolué et investissent le numérique de nouvelles potentialités (Chevret-Castellani et Labelle, 2018).

Il ne s'agit plus en effet d'associer « humanités » et « numérique », mais par le numérique d'allier la « performance » et l'humanité. Dans sa présentation au Forum sur la Gouvernance de l'Internet (FGI) le 5 juillet 2018, Mounir Mahjoubi, Secrétaire d'État auprès du Premier ministre en charge du numérique, explique que face aux enjeux de protection de la vie privée des citoyens « ces inquiétudes ne doivent pas prendre le pas sur l'innovation mais favoriser une approche raisonnée où le numérique permet d'allier performance et humanité [...] ». L'idée est de former des développeurs en mesure de trouver des solutions à la régulation. En ce sens, Alain Assouline, président du syndicat des TPE-PME, des indépendants du numérique Cinov-IT et du réseau d'écoles de développeurs WebForce3, parle d'un « design de la régulation » qui se définit comme l'objet même de cette formation. Il préconise une régulation qui reposerait sur d'autres valeurs que celles constitutives des lois.

17. « Mise en œuvre du RGPD en France, quel État ? », Chaire Valeurs et politiques des informations personnelles, Institut Mines-Télécom, Paris, 12 décembre 2018.
18. L'État n'ayant pas renouvelé son financement pour 2011, le FDI a cessé ses activités en décembre 2010.

Ce nouveau type de régulation doit favoriser le développement de la « French Tech ». Dans le cadre de la manifestation « Mise en œuvre du RGPD, quel État ? »[19], Paula Forteza, députée de La République en marche, rapporteure du projet de loi intégrant au droit français le règlement général sur la protection des données (RGPD), a fait référence au discours sur la régulation prononcé par le Président de la République Emmanuel Macron, lors du FGI du 12 Novembre 2018. Elle a évoqué en particulier son message adressé aux développeurs, selon lequel la France doit devenir un acteur majeur de la régulation, au sens technique du terme. Elle confirme la nécessité de s'ouvrir à la société civile ; « chaque citoyen devra être régulateur de lui-même ». Dans la continuité avec les rapports publics évoquant les « Humanités Numériques », il s'agit de faire acquérir des compétences permettant à l'homme d'entretenir un certain type de rapport avec la machine, rapport donnant le pouvoir à l'individu de s'autoréguler. Paula Forteza insiste sur la nécessité d'un développement de l'innovation et reprend également l'idée du Président de la République pour défendre la « French Tech », lors de la séance plénière de débat à l'Assemblée nationale du 9 octobre 2018[20].

Il s'agit de former des citoyens qui seront en mesure de réguler eux-mêmes Internet, par leur capacité à fusionner avec la machine. L'idée n'est pas neuve et n'est pas sans rappeler la « noosphère ». Comme le soulignait à juste titre Roger Bautier en 2003, elle a pu inspirer précédemment la régulation audiovisuelle. Ainsi, alors même qu'il est nommé à la Haute Autorité de la Communication Audiovisuelle, Philippe de Broglie publie la même année un ouvrage où il prédit une synthèse noosphérique au sein de laquelle fusionneront les consciences individuelles formant un type de supra-conscience (Bautier, 2003). La manière dont nombre d'acteurs publics envisagent désormais la régulation comme permettant de parvenir à une homéostasie du corps politique semble s'inspirer de conceptions bien spécifiques en provenance des théories de la cybernétique. Les discours des acteurs – comme nous allons le voir – sont comparables aux discours d'escorte des années 1990 concernant les « autoroutes de l'information ». S'il ne s'agit plus d'envisager l'internaute comme

19. « Mise en œuvre du RGPD en France, quel État ? », Chaire Valeurs et politiques des informations personnelles, Institut Mines-Télécom, Paris, 12 décembre 2018.

20. Forteza Paula, Débat en plénière : Réforme de la French Tech et entrée en vigueur de l'Open Data par défaut, http://videos.assemblee-nationale.fr/video.6698814_5bbca4745a388.1ere-seance--questions-au-gouvernement--croissance-et-transformation-des-entreprises-vote-solenne-9-octobre-2018?timecode=2497480 (consulté le 19 janvier 2019).

un récepteur passif tel que l'envisageaient les acteurs publics pour la régulation de l'audiovisuel, l'interactivité de l'homme et de la machine devrait conduire à une forme de régulation telle que conçue par Norbert Wiener.

Une conception de la régulation numérique ancrée dans la cybernétique

Oubliée un temps, la cybernétique avait d'abord ressurgi avec le discours d'escorte des « autoroutes de l'information » (Breton, 2009, p. 10). En effet, en 1994, le vice-président des États-Unis, Al Gore, avait tenu ce discours dans l'optique de comparer Internet à d'autres types de transports, comme le chemin de fer, permettant de développer le commerce. Cette idée avait conduit au développement d'un Internet grand public, étendant ainsi son influence à d'autres sphères que les seules sphères militaires et universitaires. Cette « utopie de la communication » a toutefois impliqué d'associer la communication à l'idée de « contrôle » ou plus précisément de régulation. Comme l'indique Philippe Breton, pour Wiener la régulation est ce qui devrait permettre de réduire le désordre causé par l'homme, l'entropie. Cette dernière étant la tendance qu'a la nature à transformer ce qui est ordonné. Du point de vue de la cybernétique, les machines traitent l'information de telle manière qu'elles peuvent lutter contre ce phénomène qui peut également toucher la société. C'est la raison pour laquelle l'information doit circuler et que les échanges nécessitent d'être libérés d'entraves. Wiener – qui garde en mémoire les catastrophes de la deuxième guerre mondiale – considère que plus la société progresse, plus l'entropie recule.

Du point de vue de la cybernétique, la communication ou plus exactement la transmission de l'information, au-delà de la valeur qu'elle représenterait, régit tous les phénomènes, naturels, humains et sociaux[21]. Comme le dit Madame Arpel à une amie qui visite sa maison dans le film *Mon oncle* de Jacques Tati : « *Tout communique !* ». C'est l'ensemble du réel qui peut dès lors s'interpréter (Breton, 2009, p. 25) en termes d'information et de communication. Cette dernière

21. La réduction de la communication conçue comme transmission de l'information a plusieurs conséquences. Elle fait tout d'abord abstraction de la différence entre ce qui est de l'ordre de la nature et ce qui est construit par la culture. C'est cette indistinction que des promoteurs de l'intelligence artificielle tendent aujourd'hui à défendre. Cette indistinction permettant la valorisation de la machine a pour corollaire une re-naturalisation de l'homme. Or, les complexités du cerveau humain ne résident pas dans la capacité de calculer ou décoder, mais bien dans les ressorts de l'esprit à interpréter grâce à la pensée et aux émotions.

englobe également les problèmes de droit « en ce sens qu'il s'agit de problèmes de contrôle régulier et repérable de certaines situations critiques » (Wiener, 1950, 2014, p. 138). Dans *Cybernétique et société – L'usage humain des êtres humains*, Wiener évoque le rôle de la loi et la définit comme moyen de contrôler. Il écrit : « On peut définir la loi comme le contrôle moral appliqué aux modes de communication, tels que le langage, particulièrement lorsque ce caractère normatif se trouve contrôlé par une autorité assez solide pour donner à ses décisions une sanction sociale efficace. » (Wiener, 1950, 2014, p. 133). Loin d'une conception rousseauiste selon laquelle elle est l'expression de la volonté générale (conception reformulée dans la tradition juridique française par Carré de Malberg), la loi est ici conçue comme le cadre de l'exercice de l'autorité suivant une vision pragmatique, voire utilitariste, de l'efficacité.

Une des conséquences d'une conception de la communication comme transmission de l'information impliquant une régulation comparable à une homéostasie est d'envisager, comme nous l'avons souligné, la loi comme le cadre de l'exercice de l'autorité. Or, l'abandon d'une compréhension de la loi comme expression de la volonté générale a plusieurs significations et implications. Il pèse sur la manière dont le droit de l'internet se construit. Cette construction repose sur deux tendances dont l'idée d'autorégulation du citoyen – sa capacité à être régulateur de lui-même –, n'est finalement que l'aboutissement : l'influence des sciences informatiques d'une part et celle d'une pensée économique anglo-saxonne qui conduit à la patrimonialisation des contenus numériques, d'autre part.

L'enchâssement des discours publics et politiques dans le droit relatif à l'internet : prédominance du vocabulaire informatique et tendance à la patrimonialisation individuelle des contenus numériques

Les discours publics et politiques sur l'internet ne renvoient plus, en France et de manière plus générale en Europe, à un modèle calqué sur celui de la télévision, modèle qui a longtemps prévalu faute de compréhension des caractéristiques du média. Il supposait un internaute passif, victime des contenus circulant sur le web (pornographie, cybercriminalité, etc.). Ces discours influencent davantage aujourd'hui la manière dont le droit se construit à la fois par la prédominance d'un vocabulaire véhiculé par les informaticiens et par une tendance à concevoir les contenus numériques comme relatifs à un patrimoine individuel.

La prédominance du vocabulaire informatique

Les discours publics et politiques s'inspirent d'abord de ce que les informaticiens nomment la « donnée », c'est-à-dire l'information, et vont déterminer la manière dont se construit le cadre juridique. Ce dernier, en France et en Europe, a d'abord retenu l'idée de « données nominatives ». La notion de données à caractère personnel (ou données nominatives) est développée dans la loi du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, dans la Convention 108 du Conseil de l'Europe de 1981 et son protocole additionnel de 2001, puis dans la directive européenne de 1995[22]. Le cadre juridique le plus récent se rapporte au RGPD du 27 avril 2016, entré en vigueur le 25 mai 2018, et à la loi sur la protection des données personnelles du 21 juin 2018.

Le terme de « données », notamment mobilisé aujourd'hui dans le champ de l'informatique, renvoie à des indications enregistrées par les machines qui vont permettre une analyse ou une recherche automatique d'informations, à l'instar d'une « base de données », de « données documentaires » ou de « données lexicales ». Toutefois, son origine est d'abord mathématique. Il n'est d'ailleurs pas accessoire de rappeler que Cournot – qui en est le premier promoteur – a été mathématicien, mais aussi économiste et inspecteur général de l'instruction publique. Dans son *Essai sur les fondements de nos connaissances*, il l'entend comme ce qui est formulé dans l'énoncé d'un problème : « [la langue de l'algèbre] fournit les moyens de soumettre les grandeurs aux mêmes opérations de calcul, sans distinction de données et d'*inconnues.* » (Cournot, 1851, p. 391). Dans le texte de Cournot, le terme de « données » n'est pas seulement utilisé pour les mathématiques, mais aussi pour les sciences de la nature avec des « données expérimentales », « données d'observation », ou plus généralement « données de la raison ». Aujourd'hui, les « données » font l'objet d'une « gestion », que ce soit dans le cadre de la mise en œuvre du RGPD par les entreprises, dans celui de l'ouverture des données (Open Data), jusqu'aux préconisations de l'Agence Nationale de la Recherche qui exige la présentation d'un « plan de gestion des données » dans les six mois après le démarrage de projets financés (https://anr.fr/fr/lanr-et-la-recherche/engagements-et-valeurs/la-science-ouverte/, consulté le 13 août 2019).

22. Les intitulés exacts des deux textes européens qui ont précédé le RGPD sont : Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel du 28 janvier 1981 et Directive 95/46/CE du Parlement européen et du Conseil, du 24 octobre 1995, relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données.

La tendance du droit de l'internet à la patrimonialisation individuelle des contenus numériques

L'autre conception véhiculée dans les discours de politiques publiques s'ancre, surtout depuis 2004, dans un point de vue supposant les contenus numériques produits relativement aux données personnelles. Cette vision implique de plus en plus de circonscrire les enjeux relatifs à l'internet à la question de la « gestion » de ses propres données. C'est l'esprit qui présidait à la loi pour la confiance en l'économie numérique (LCEN) du 21 juin 2004 mais encore à celle qui lui a succédé, la loi pour « une République numérique » du 7 octobre 2016. Dans cette loi dite « Lemaire » de 2016, l'article 63 « mort numérique » introduit des directives concernant la « gestion » *post-mortem* de ses données à des ayants droit. La patrimonialisation individuelle des données a dominé un débat orchestré par le Conseil supérieur du notariat dans le cadre de la co-écriture de la loi (Alexis *et al.*, 2017) et cette idée a influé sur le texte définitif. Il a été considéré que les données numériques pouvaient être définies comme des biens transmissibles aux héritiers. Cette conception économique s'éloigne d'autres conceptions, personnaliste ou communaliste. En effet, le personnalisme dont la tradition juridique française a été porteuse jusqu'au XXI^{e} siècle – et que véhicule encore la CNIL – va s'attacher aux droits attribués à une personne parce qu'elle mérite ontologiquement le respect en tant qu'être humain. Le communalisme va, de manière distincte, amener à considérer que les données sont constitutives de biens communs et ne peuvent donc pas faire l'objet d'une appropriation. Aujourd'hui, celui-ci a pour mérite de souligner que la réduction de la question des « données » à des enjeux de « gestion » par l'individu néglige la dimension collective. Cette question est reprise par les Sciences de l'information et de la communication (Sic) qui amènent à prendre en compte cette dimension collective dans laquelle s'ancre l'info-communicationnel. Face à ces tendances dans la manière de régir le numérique, les Sic sont en effet en mesure de proposer d'autres options, notamment à travers la notion de « trace », qu'elles contribuent à construire.

Conclusion

Les « Humanités Numériques » s'inscrivent aujourd'hui dans des conceptions systémiques de l'environnement numérique nourries par certains acteurs publics et politiques. En effet, elles représentent un contexte politique en soi dans la mesure où elles ne renvoient plus seulement à des méthodes portées par la recherche ou même à un type d'enseignement lié à l'informatique ; il s'agit de former des citoyens de demain qui soient capables d'une part de développer des systèmes informatiques, de promouvoir une capacité d'innovation et de croissance économique de la « French Tech », et d'autre part de s'autoréguler. Les discours publics et politiques sur la « société numérique » dans lesquels elles sont intégrées ne sont pas sans rappeler les discours d'escorte concernant les « autoroutes de l'information », dans les années 1990. C'est la même influence de la cybernétique de Wiener qui semble s'y exercer.

Cette approche systémique englobe en effet également les capacités qu'auront les citoyens de devenir régulateurs d'eux-mêmes et les développeurs d'intégrer les moyens techniques de la régulation, à l'instar de ce qu'ils font aujourd'hui par exemple en *Privacy by design.* Elle induit d'envisager une homéostasie du corps politique. Notons que Wiener lui-même contestait l'idée même de cette homéostasie. Ainsi, il considérait comme une « croyance », l'idée d'une régulation par la libre concurrence du marché telle qu'elle se développait aux États-Unis déjà à son époque[23]. Cependant, Wiener tendait aussi à promouvoir une régulation ou plus exactement une autorégulation jusqu'à espérer une délégation de la décision politique au profit des machines (Breton, 2009, p. 36). Cette promotion de la régulation – voire de l'autorégulation – par les pouvoirs publics, comme alternative à la loi et à la réglementation, représente sinon une exclusivité pour le moins une spécificité. Les autres médias ne sont pas à ce point impactés par des modifications de conception de normes qui doivent les régir. Nous entendons peu parler de « régulation de la presse », mais bien encore essentiellement de « droit de la presse ».

Cette évolution des conceptions de la régulation d'Internet semble comme le point d'aboutissement de deux tendances, celle déjà

23. L'homéostasie du corps politique est envisagée ainsi par Wiener : « Selon une croyance, courante dans de nombreux pays et élevée au rang d'un article de foi officiel aux États-Unis, la libre concurrence serait elle-même un processus homéostatique : dans un marché libre, l'égoïsme individuel des hommes d'affaires cherchant à vendre aussi cher que possible et à acheter au prix le plus bas résulterait en une dynamique des prix stable, favorisant le plus grand bien commun » (Wiener, 1948, 2014, p. 191).

présente depuis les années 1970 qui consiste à reprendre le vocabulaire de l'informatique, comme par exemple le terme de « données », pour construire le droit et une autre qui implique, surtout depuis le début du XXIe siècle, la patrimonialisation individuelle des contenus numériques. Face à ces tendances, les Sic sont en mesure de relever le défi permettant de trouver des alternatives, d'autres normes pouvant réguler le média, notamment en développant le concept de « trace ». En effet, concernant le numérique, elles vont envisager à la fois une inscription de signes sur un support et une relation dans le cadre d'une interaction ; elles critiquent aussi le traitement algorithmique effectué par les dispositifs numériques (Rieder, 2010). Elles le font de manière différente de celle des sciences informatiques. En effet, lorsque les concepteurs de systèmes informatiques nomment « traces » des productions d'individus, ces concepteurs exploitent en fait des situations de communication. Les informaticiens utilisent le terme de trace, mais dans deux sens différents de celui donné par les Sic : c'est-à-dire soit la valeur interne à la mémoire qui rend explicite le déroulement d'un programme, soit ce qui se rapporte à la production d'un utilisateur d'un logiciel. Les conceptions informatiques de la « donnée » influent aujourd'hui sur la construction du droit relatif à l'internet ; elles déterminent également des définitions particulières de la « trace » éloignées de celles transmises par les Sic.

La « trace » est une notion en Sic qui permet d'englober différents phénomènes observés. Elle ne se rapporte pas à un concept (Jeanneret, 2011), mais elle ne se construit pas non plus comme une notion floue. Il s'agit plutôt d'une notion souple permettant de circonscrire des phénomènes étudiés aussi bien, comme le souligne Sylvie Leleu-Merviel, d'un point de vue info-communicationnel que d'un point de vue info-documentaire (Leleu-Merviel, 2013). L'article « Traces » de Camille Alloing du *Publictionnaire* apporte sur ce point une synthèse éclairante de la manière dont les Sic se sont emparées de cette notion et quel avenir elles peuvent lui offrir. Il met notamment en évidence les apports de Louise Merzeau, d'Olivier Ertzscheid et de Fanny Georges concernant la manière dont les traces sont constitutives des identités numériques des individus. Sans prétendre ici à l'exhaustivité de leur recensement, les projets de construction du concept de trace numérique sont nombreux. Ils ont été pensés, avec Louise Merzeau, par la médiologie. Ils sont portés par la sémiologie d'Yves Jeanneret et par une lignée proche – comme celle de Béatrice Galinon-Mélénec avec le concept de « signe-trace » –, ou encore par la « tracéologie » élaborée par Linda Idjéraoui-Ravez et Nicolas Pélissier.

Ils rassemblent aussi des chercheurs de ces différents horizons, tel que « Human Trace » soutenu par l'Unesco[24]. Au-delà de la pluralité des approches, la vivacité des analyses détermine aujourd'hui des enjeux qui permettront que la trace soit pensée autrement que comme un objet saisi par un robot ou une opération effectuée par des logiciels.

POSITIONNEMENT

Les discours publics et politiques ont aujourd'hui tendance à véhiculer une conception de l'enseignement en « Humanités Numériques » et informatique comme liée à une manière non seulement de construire des enseignements, d'instruire et de former, mais aussi de projeter et de promouvoir une certaine conception de la « société numérique ». Pour certains acteurs publics et politiques, il en découle un type de citoyenneté impliquant la capacité à intervenir sur Internet et à en modifier la configuration, ce qui suppose une formation au métier de développeur. Ces spécialistes de l'informatique participeraient à la fois au déploiement économique de la « French Tech » et à l'invention de systèmes intégrant la régulation. Cette conception de la régulation trouve un de ses points d'ancrage dans la cybernétique de Norbert Wiener dont les fondements avaient déjà ressurgi, en 1994, dans les discours sur « les autoroutes de l'information » du vice-président états-unien Al Gore. Elle semble comme le point d'aboutissement d'une évolution de la façon de penser la régulation ayant présidé à la construction du droit de l'internet et s'accompagne de la diffusion du langage informatique, à l'instar du terme de « données », s'orientant vers une patrimonialisation individuelle des contenus du web. Cet aboutissement n'est cependant pas inéluctable. Les Sic sont en effet en mesure de proposer de nouveaux concepts permettant de réorienter cette manière d'envisager les normes qui doivent régir l'internet, notamment au regard de la notion de « trace ».

BIBLIOGRAPHIE

Alexis L., Chevret-Castellani C., Labelle S. et Mouratidou E., 2016. « Vers le partage du *savoir-écrire* la loi ? Analyse sémio-discursive du dispositif republique-numerique.fr ». *Semen*, n° 42, pp. 93-115.

Alloing C., 2018. « Traces numériques ». *Publictionnaire. Dictionnaire encyclopédique et critique des publics*. Mis en ligne le 09 janvier 2018, http://publictionnaire.humanum.fr/notice/traces-numeriques/. Consulté le 19 janvier 2019.

24. Human Trace, https://onhumantrace.hypotheses.org/3968 (consulté le 2 mai 2019).

Bautier R., 2003. L'internet comme cerveau mondial. *Esprit critique*, Automne 2003, vol. 5, n° 4, http://www.espritcritique.org. Consulté le 2 mai 2019.

Bouquillion P. et Matthews J.-T., 2010. *Le Web collaboratif. Mutations des industries de la culture et de la communication*. Grenoble, Presses universitaires de Grenoble, 150 p.

Breton P., 2009. *L'utopie de la communication – Le mythe du village planétaire*, Paris, La découverte, 171 p.

Chevret-Castellani C. et Labelle S., 2018. L'injonction à l'innovation dans la régulation des algorithmes : créativité et idéologie dans les politiques publiques, Actes du Congrès de la SFSIC « Création, créativité, médiations », vol. 1, Saint-Denis, pp. 104-114.

Cournot A.-A., 1851. *Essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique*, Wikisource, https://fr.wikisource.org/wiki/Livre:Cournot_-_Essai_sur_les_fondements_de_nos_connaissances.djvu. Consulté le 19 janvier 2019.

Dulong de Rosnay M., 2016. *Les golems du numérique*. *Droit d'auteur et Lex Electronica*. Paris, Presses des Mines, 260 p.

Jeanneret Y., 2011. « Complexité de la notion de trace. De la traque au tracé » dans B. Galinon-Mélénec (dir.), *L'homme trace. Perspectives anthropologiques des traces contemporaines*, Paris, CNRS Éditions, pp. 59-86.

Labelle S., 2001. « Société de l'information, à décrypter ! » dans *Communication et langages*, n° 128, pp. 65-79.

Leleu-Merviel S., 2013. « Traces, information et construits de sens. Déploiement de la trace visuelle de la rétention indicielle à l'écriture ». *Revue de l'Association pour la Recherche Cognitive*, n° 59, pp. 65-88.

Mounier P., 2018. *Les Humanités Numériques – Une histoire critique*. Paris, Maison des Sciences de l'homme, 180 p.

Rieder B., 2010. Pratiques informationnelles et analyse des traces numériques : de la représentation à l'intervention. *Études de communication*, 35, p. 91-104, http://journals.openedition.org/edc/2249. Consulté le 19 janvier 2019.

Triclot M., 2006. « Les implications politiques de la cybernétique américaine ». *Araben*, n°3 (janvier), pp. 54-63.

Valluy J., 2018. « Segment – Deuxième définition (2010/2017…) : technique, éditoriale, a-disciplinaire en humanités classiques » dans *Transformations des États démocratiques industrialisés – TEDI* – Version du 4 février 2019, identifiant de la publication au format Web : http://

www.hnp.terra-hn-editions.org/TEDI/article434.html. Consulté le 14 août 2019.

Valluy J., 2015. « La face cachée des Humanités Numériques : les convergences techniques et théoriques masquent l'ampleur des différenciations disciplinaires ». Colloque CJC-SC, 3/5 juin 2015, http://www.hnp.terra-hn-editions.org/TEDI/IMG/pdf/jv_cjcsc_050856b.pdf. Consulté le 19 janvier 2019.

Wiener N., 2014 [1950]. *Cybernétique et société. L'usage humain des êtres humains*, Paris, Éditions du Seuil, 224 p.

Wiener, N., 2014 [1948]. *La cybernétique* : *information et régulation dans le vivant et la machine*, Paris, Éditions du Seuil, 376 p.

CHAPITRE V

Questionnement éthique par rapport au numérique

Jean-Claude Domenget
Université Bourgogne Franche-Comté, Elliadd EA 4661
Carsten Wilhelm
Université de Haute-Alsace, CRESAT UR 3436
Cécile Dolbeau-Bandin
Université de Caen, CERREV EA 3918

Résumé : L'intérêt d'un questionnement éthique en contexte info-communicationnel, en lien avec les stratégies et les pratiques de recherche, s'accroît à la mesure qu'outils, méthodes, objets et valorisation intègrent le numérique. L'importance croissante des transformations opérées par ce biais nécessite à la fois un positionnement éthique de la part de la société et des chercheurs en Sic ainsi qu'une prise en compte de ces transformations dans les méthodes de construction des projets de recherche. Notre récente enquête menée auprès des chercheurs en Sic et une mise en perspective des enjeux pour les Humanités Numériques permettent de revenir sur l'expérience des Sic françaises et leurs interrogations face à ces défis.

Mots-clés : éthique de la recherche, numérique, Sic, Humanités Numériques, dynamique internationale.

Abstract: The interest for ethical approaches towards info-communicational research strategies and practices grows, as tools, methods, objects and science communication itself integrate digital technology. The increasing importance of these transformations requires both an ethical positioning on behalf of society and the researchers in information and communication sciences (Sic) as well as taking these transformations into account in the methods of constructing research projects. Our recent survey of Sic researchers and a perspective on the challenges for digital humanities allow us to review the experience of French Sic and their questions in the face of these challenges.

Keywords: research ethics, digital, information and communication sciences, digital humanities, international dynamics.

L'objet de ce chapitre est de mettre en lumière le nécessaire questionnement éthique qui est de plus en plus souvent invoqué de nos jours dans les métiers de la recherche comme dans la société en général. Traduisant une montée des valeurs de transparence et de responsabilité, la convocation de l'éthique dans les relations entre la recherche et le numérique donne lieu à la mobilisation d'anciens codes déontologiques, une attention renouvelée pour l'intégrité scientifique ou encore une volonté de régulation juridique dans la collecte et le traitement massif des données personnelles. Ces quelques exemples illustrent l'intérêt d'un questionnement éthique en contexte info-communicationnel numérique, en lien avec les logiques d'acteurs, les stratégies et les pratiques de recherche ainsi que l'édition de documents (Balicco, Broudoux, Chartron, Clavier, Paillart, 2018).

Sans qu'il soit toujours bien défini, le questionnement éthique rappelle que le numérique présente de nombreuses caractéristiques qui le rendent capable de nous faire agir autrement (Boullier, 2019) : la persistance des traces, la duplication des contenus (Boyd, 2008), l'importance du calcul ou encore l'introduction de l'incertitude. Ainsi, le numérique, en tant qu'objet et condition de la recherche pose également des questions éthiques nouvelles à plusieurs égards (Vitali-Rosati, 2012) dans les transformations des interactions entre l'humain et la machine, contribuant à redéfinir le contrat moral avec la société (Mounier, 2018).

Ces problématiques que nous choisissons d'appeler « éthiques » interrogent les chercheurs mais plus largement la société sur la place du calcul, des machines, des objets dits « intelligents » dans le cadre du travail, du loisir ou plus largement des activités produisant des données de recherche comme des activités de communication quotidiennes. Il est ici question de la responsabilité des chercheurs concernant les traces maniées comme objets et comme résultats

de leurs recherches, la provenance des données et la connaissance des processus les produisant. Dans ce contexte, les Humanités Numériques, s'appropriant les méthodes digitales pour l'étude de corpus variés, incluant des productions intellectuelles et personnelles manifestes dans des traces numériques, s'exposent alors à une double injonction éthique. D'une part, le passage par le calcul nécessite une documentation méticuleuse des contextes de production et d'analyse des données. D'autre part, la présence de données personnelles dans ces traces demande une attention particulière au respect des droits individuels et ce depuis l'avant-RGPD (règlement général sur la protection des données).

L'ampleur de ces problématiques a conduit à la création en 2018 du Groupe Éthique et Numérique en Information-Communication (GENIC)[1] qui vise à développer une réflexion compréhensive des enjeux éthiques du numérique. Il organise sa réflexion à travers plusieurs groupes de travail. Un premier groupe s'est donné la tâche de rassembler les définitions courantes autour de l'éthique de la recherche afin de pouvoir suivre les évolutions de ce terrain en mouvement à travers les discours qui l'accompagnent et les conventions qui se créent, qu'il s'agisse de la bioéthique, bien visible dans le discours médiatique actuel, ou de champs plus liés à l'information-communication comme la participation en ligne. Un deuxième groupe étudie l'évolution des réglementations qui s'appliquent en contexte numérique, ayant comme but de protéger les individus dans une visée éthique, en particulier au regard des données sensibles. Au sein de ces nouvelles normes, se posent alors certaines questions : quelles sont les délégations de pouvoir qui s'opèrent ? Quelles responsabilités sont prévues pour quels acteurs (CNIL, Comité consultatif national d'éthique…) ? Un troisième groupe de travail s'applique à dresser un bilan comparatif des travaux, réglementations et définitions sur le plan international et dans une visée interdisciplinaire, tout en privilégiant les sources équivalentes aux Sic françaises. Il s'agit de sonder non seulement les résultats mais également les processus de réflexion et d'écriture collective qui font naître les documents dans leurs contextes socioculturels respectifs. Un quatrième groupe s'est donné pour tâche d'étudier un corpus de publications mobilisant les données numériques et/ou des questionnements éthiques dans les revues du champ de l'information-communication. Un premier objectif est d'observer la part accordée

1. Le groupe GENIC est né sous l'égide de la SFSIC, à l'initiative de deux de ses administrateurs, Jean-Claude Domenget, en tant que vice-président recherche et Carsten Wilhelm, en tant que vice-président relations internationales, pour le mandat 2018-2020.

à la méthodologie (modalités de collecte et de traitement, types de données, etc.) et au questionnement éthique. Un second est de repérer quelles dimensions de l'éthique sont abordées. Les productions de ces quatre groupes seront à valoriser, ce qui constitue l'objet d'un dernier groupe de travail. Des publications, en commençant par ce chapitre, viendront nourrir la discussion de la communauté. En plus des moyens classiques de valorisation scientifique, une mise à disposition de fiches pratiques est envisagée ainsi que l'ouverture d'un espace de dialogue pour les chercheuses et les chercheurs de notre discipline.

Ainsi, nous avons pensé ce chapitre autour de l'idée que les transformations précédemment évoquées questionnent le positionnement éthique du chercheur en Sic dans la mise en place de ses projets de recherche. Dans un premier temps, nous présenterons le questionnement éthique qu'il s'agit de mener à chaque étape d'une recherche portant sur ou utilisant des données numériques. Dans un second temps, nous proposerons une réflexion plus large, intégrant l'importance de participer à la dynamique internationale et à l'ouverture socioculturelle autour de ces questions, pour conclure sur les usages et les interprétations des méthodes dites digitales, au cœur des Humanités Numériques.

Éthique – les étapes de la recherche en contexte numérique

Le numérique conduit à une attention renouvelée au sujet des relations entre éthique et recherche. Au-delà de l'intégrité scientifique et de ses sempiternelles interrogations sur les modalités de prévention de conduites proscrites dans la profession, telles que le plagiat ou la publication de faux résultats, travailler sur des terrains numériques pose des questions éthiques. À titre d'exemple, la participation de chercheurs à un groupe de discussion en ligne conduit à s'interroger sur le positionnement transparent ou non de ses membres, question héritée de l'anthropologie et de l'observation participante. Lorsque l'on souhaite analyser des sujets polémiques ou donnant lieu à des controverses scientifiques, faut-il se présenter, se déclarer en tant que chercheuse, chercheur ou cacher cette facette de son identité afin d'avoir accès à certaines conversations ? La réponse est rarement binaire et la prise en compte d'éléments de contexte s'avère être indispensable. Les Humanités Numériques, en adoptant plus largement des corpus numérisés et des méthodes d'analyse informatisées s'exposent d'autant plus à ces questions que l'attrait pour les méthodes précède généralement leur analyse critique et

que le mariage interdisciplinaire, nécessaire au fonctionnement de tels projets, demande souvent le temps long de la négociation des objectifs et des procédures.

Réfléchir aux transformations du questionnement éthique tout au long d'un projet de recherche en contexte numérique va bien au-delà de l'objectif de cette partie. Nous ne pouvons esquisser ici que quelques dimensions de la réflexion nécessaire et nous renvoyons les personnes intéressées à deux travaux réflexifs. Le premier porte sur une proposition de cadrage conceptuel d'une réflexion éthique pour conduire des recherches en régime numérique (Simonnot, 2018), lequel aborde l'éthique en recherche autour de la question de la responsabilité ; les transformations du numérique dans les activités de recherche touchant notamment la publication des résultats, la confidentialité des données et le consentement éclairé des participants ; ainsi que la montée d'une institutionnalisation des procédures éthiques dans de nombreuses disciplines en sciences humaines et sociales. Le second travail est issu d'une enquête sur l'éthique des pratiques de recherche liées au numérique en Sic, (Domenget, Wilhelm, 2018)[2], laquelle a souhaité mettre en débat le questionnement éthique en régime numérique, en saisissant le positionnement des chercheuses et des chercheurs en Sic vis-à-vis des principes éthiques dans leurs pratiques de recherche, en traitant des problématiques spécifiques aux données numériques de la recherche, en demandant leur avis et en recueillant leur expérience concernant les pratiques de validation éthique de la recherche, etc. Un second objectif de cette enquête était d'évaluer le rôle que pouvait jouer une société savante comme la SFSIC dans la réflexion sur ces questions, en gardant bien à l'esprit la nécessaire ouverture internationale et l'ambition transculturelle (que nous aborderons dans la deuxième partie de ce chapitre) qui permet de mieux situer le contexte de la recherche en France.

Fort de ce rappel, nous avons choisi de focaliser notre réflexion sur quatre étapes de la recherche que sont : la définition de la question de recherche ; le choix méthodologique ; le terrain proprement dit à travers le recueil, l'analyse et la préservation des données numériques collectées ; ainsi que la publication et le partage des résultats. Ce découpage permet de souligner non seulement que le

2. Cette enquête a comporté une partie questionnaire diffusée auprès des membres de la SFSIC (lequel a recueilli 119 réponses) et une partie entretiens (15 entretiens) avec des collègues ayant déjà dû gérer ces enjeux éthiques dans leurs projets de recherche ou ayant une réflexion avancée sur le sujet.

questionnement éthique ne peut se cantonner à une phase amont de la recherche, lors du dépôt du projet, mais aussi que l'amplification et les transformations du questionnement induites par le numérique ne peuvent être isolées des changements du contexte de la recherche (changements législatifs, montée d'une institutionnalisation des procédures de régulation éthique de la recherche, évolution de la représentation du rôle social de la recherche scientifique, etc.).

Éthique – la définition des questions de recherches

Dans une démarche de recherche scientifique, un premier aspect de l'interrogation éthique aborde les questions de recherche sur lesquelles le chercheur est amené à réfléchir. Il s'agit de s'interroger sur :

1. l'objet de recherche, indépendamment des contraintes législatives, financières, etc. faisant écho au principe d'autonomie du chercheur, très valorisé lors de notre enquête ;
2. les modalités de réalisation de la recherche, dans une approche très pragmatique et située ;
3. les objectifs de la recherche, traitant de la responsabilité sociétale de la recherche scientifique.

En effet, si l'orientation de la recherche est indéniable, par l'intermédiaire des financements des programmes de recherche, lesquels imposent aujourd'hui un questionnement éthique, le choix d'un objet de recherche et de la manière de l'interroger pose principalement la question de la responsabilité du chercheur. « À mon sens l'éthique c'est avant tout une question de responsabilité, de la responsabilité de ce qu'on fait » (ITW 6)[3]. Or, l'assise de la responsabilité et l'*agentivité* morale, traditionnellement situées au niveau de l'individu sont de plus en plus perçues comme relationnelles et impliquent un réseau d'acteurs (Ess, 2002). La responsabilité d'un chercheur[4] ne peut être questionnée indépendamment des interactions avec les différentes parties impliquées (financeurs, population d'études, etc.). Ainsi, évoquant une recherche impliquant

3. Nous reproduisons ici des extraits d'entretiens issus de l'enquête sur l'éthique des pratiques de recherche liées au numérique en Sic (Domenget, Wilhelm, 2018). 15 entretiens ont été réalisés de février à juin 2018 auprès d'enseignants-chercheurs, professeurs et maîtres de conférences, un ingénieur de recherche et un praticien de la communication.

4. L'utilisation du genre masculin est utilisée afin de faciliter la lecture et n'a aucune intention discriminatoire.

des personnes ayant une maladie grave, une chercheuse explique que le questionnement éthique « c'est essentiel parce que cela me permet de respecter mes valeurs morales. Ça m'aide à accepter ce que je fais. Au niveau moral, cela donne presque une légitimité en tant que chercheur » (ITW 1).

Cette dimension de la responsabilité sociétale des chercheurs renvoie à la question de la *capacité* d'agir des acteurs. Dans ce but, S. Szoniecky propose une approche très située, pragmatique de l'éthique en recherche[5], en s'appuyant sur les travaux de Deleuze et de Spinoza ou encore la réflexion éthique de Morin autour de la méthode. Il propose en outre de réfléchir à une modélisation du pouvoir d'agir de l'Internet des Objets (IOD ou objets connectés) qui a pour objectif d'évaluer et de comparer l'éthique de ces technologies afin de comprendre leurs impacts sur nos vies quotidiennes (Szoniecky, Safin, 2017). Ce questionnement est essentiel pour dépasser par exemple les réglementations autour de l'éthique des données ou mener des recherches sur des thématiques susceptibles d'être proscrites par les instances de régulation (à l'image des conduites à risque voire suicidaires chez certains jeunes sur les médias socionumériques). En effet, s'interroger sur la *capacité* d'agir d'un chercheur, c'est dépasser une approche purement législative ou normative, voulant définir ce qu'un chercheur aurait le droit de faire ou de ne pas faire et s'intéresser davantage aux conditions de réalisation de ses recherches. Est-ce qu'il dispose des ressources nécessaires pour, par exemple, répondre aux exigences de sécurisation des données collectées lors d'une enquête ? Et c'est en fonction de la situation décrite que la question éthique se pose : « je fais ou je ne fais pas » (ITW 8).

Éthique – les choix méthodologiques

Le deuxième aspect du questionnement éthique concerne les choix méthodologiques. « Si on a affaire, et moi j'ai constamment affaire à des données sur lesquelles j'estime qu'il faut savoir placer une éthique, même au-delà des garde-fous techniques, juridiques, des normes, etc. Je pense que ça serait intéressant de réfléchir à nos approches méthodologiques » (ITW 3). Ainsi, le questionnement éthique lié au travail méthodologique est perçu par plusieurs collègues interviewés comme une source d'amélioration de la recherche. « Ça m'oblige à être plus rigoureuse, dans la manière de réfléchir à ma méthodologie en amont. Parce que le fait de savoir que ça va être compliqué, cela

5. Nous nous référons ici aux échanges ayant eu lieu lors des réunions du groupe GENIC.

nécessite de prendre en amont ce facteur-là » (ITW 1). « Avec l'idée, être plus éthique, c'est aussi faire une meilleure recherche, j'en suis persuadé en SHS. Prendre soin de ses enquêtés, prendre soin de ses données, prendre soin de son terrain, c'est le meilleur moyen de produire des choses intéressantes » (ITW 3).

Cette importance des choix méthodologiques vient compléter les thématiques déjà évoquées de la transparence du chercheur ou du consentement éclairé des participants qui sont le plus souvent abordées. En effet, l'évolution des méthodologies de recherche liées à des données numériques conduit à orienter le débat sur le possible transfert des codes et des principes éthiques des méthodes traditionnelles de nature qualitative – observation non participante, entretien en profondeur, etc. – aux méthodes dites « numérisées » – *digital methods*, ethnographie en ligne (Théviot, 2013) –, cela d'autant plus avec le développement de méthodes mixtes (« *mixed methods research* ») dépassant l'opposition entre méthodes traditionnelles et méthodes « numérisées » (Iphofen, 2017 ; Woodfield, 2017). Avec l'idée que dans ce type de transfert, le travail reste une affaire de bricolage. « Il n'y a pas de textes de cadrage justement… y compris chez les anthropologues. C'est à toi à faire ta propre éthique, ce qui te paraît acceptable, on fait des choix… Les textes, ils vont te dire de préserver l'anonymat des personnes quand tu traites de données sensibles, des choses comme cela mais ils ne vont pas te dire comment tu dois le faire » (ITW 2).

Éthique – le terrain : autour des données de la recherche

Le troisième aspect concerne le recueil, l'analyse et la préservation des données collectées. Liée aux choix méthodologiques, la réflexion éthique portant sur le terrain de recherche a conduit à évoquer les difficultés rencontrées lorsque plusieurs chercheurs ont voulu par exemple faire valider leur questionnaire de recherche par la commission éthique de leur université. Les reproches portent notamment sur la longueur des délais de réponse, l'aspect chronophage et compliqué de la procédure. En fait, la méconnaissance de cette dernière conduit à un sentiment de blocage administratif voire d'une interprétation trop rigoriste des règles édictées dans les chartes éthiques de la recherche scientifique. « On a voulu faire passer à la commission d'éthique de l'université notre questionnaire. Ça a été extrêmement long […]. C'était très compliqué. […] C'était bloqué pour des questions purement administratives en fait. On ne connaissait pas les procédures » (ITW 1). Plus largement, l'abondance, l'accès facilité

au recueil des données tend à faire oublier que l'une des premières questions à se poser devrait être : « est-ce que je collecte des données qui ont une pertinence pour ma recherche ? » (ITW 3).

En lien avec cette interrogation, le questionnement éthique invite à éviter toute quantophrénie des données, laquelle risque d'éloigner les chercheurs du terrain et de ses clefs d'interprétation. En effet, « L'abondance soudaine des données, en effet, risque toujours de susciter une sorte d'ivresse, signant le triomphe d'un nouveau positivisme allant chercher dans les seules données les questions et les réponses, ce qui fixerait l'agenda de la recherche » (Wieviorka, 2013, p. 43). La possibilité aujourd'hui de recueillir et d'exploiter une grande quantité de données ne doit en effet pas conduire à négliger, par facilité, le travail essentiel de problématisation et de questionnement du projet de recherche, tel que rappelé précédemment. D'autant plus, que ce questionnement permettrait de distinguer le travail d'analyse scientifique des données, de celui effectué par des plateformes, dont le fonctionnement est d'agréger un maximum de données dans un seul but de performance et de marchandisation de ces dernières. Les modalités d'analyse des données posent en fait la question du contexte de production puis d'interprétation des données, rendant nécessaire une conciliation entre disciplines. Il est vrai que dans le cadre de projets interdisciplinaires, « les approches peuvent être assez antagonistes quoi… entre notamment informaticiens et chercheurs en SHS » (ITW 7).

Complétant la nécessaire réflexion sur les modalités d'analyse des données numériques, celle sur leur archivage est essentielle car l'analyse se fait en contexte. Les collègues interrogés soulignent les limites de l'injonction au partage et à l'ouverture des données de la recherche hors contexte. Au contraire, « il faut faire un plan d'archivage des données qui permette bien de restituer le contexte dans lequel les données ont été prélevées. Parce que si on analyse des données sans leur contexte, […] ça ne pardonne pas. […] On fait des analyses qui sont situées, qui sont hyper-contextualisées » (ITW 2). De même, cette réflexion sur les conditions de production, d'analyse et de stockage des données aborde les milieux numériques dans lesquels les données sont recueillies et analysées. Des données « propriétaires » des plateformes du web social sont inaccessibles aux chercheurs, non affiliés à ces acteurs et l'analyse se fait *via* des outils dont on ne connaît pas forcément le code. « Et puis, je dirais qu'au-delà de ça, il y a aussi un enjeu épistémologique, parce qu'effectivement, il y a des logiciels qui vont fonctionner avec du clic interface, du clic bouton,

mais la problématique c'est que ça va conduire les chercheurs à un modèle épistémologique qu'ils ne sont pas en mesure de maîtriser totalement, à la différence des données qu'ils gèrent directement, en programmant eux-mêmes » (ITW5).

Il s'agit ainsi d'adopter un point de vue critique concernant le travail autour des données de la recherche et l'adoption des méthodes digitales en SHS, car « si le numérique présente de nombreux avantages pour les SHS, étant donné l'ancrage dans le social ainsi que les transformations des conditions de recherche et de leur publicisation qu'il permet, le chercheur doit résister à la fascination qu'il peut exercer en faisant preuve de vigilance critique et épistémologique ; les Humanités Numériques n'étant pas au service de la technique mais de la recherche, de la société et du savoir » (Bourdeloie, 2013). Cette dimension critique incite à élargir la réflexion sur les modalités de publication des résultats de la recherche.

Éthique – la publication des résultats

Le quatrième et dernier aspect concerne le choix opéré par le chercheur dans la manière de publier ses résultats. Le questionnement éthique invite de nouveau à s'intéresser à l'anonymisation des données lors de cette phase de publication et non plus uniquement lors de la phase de terrain. En effet, l'environnement numérique et les technologiques de recherche d'information disponibles rendent nécessaires de nouvelles solutions d'anonymisation. Ainsi, dans le cas par exemple de la publication des résultats d'une analyse des interactions en ligne, un risque déjà pointé est la mise en visibilité de données personnelles ou sensibles du fait même du travail du chercheur (Latzko-Toth, Pastinelli, 2013). « Nous avons disposé d'une base de données [qui] n'a pu nous être transmise qu'après avoir été totalement anonymée. […] c'est-à-dire que toutes les identités des signataires, des porteurs de pétitions, etc. avaient été codées. Mais on s'est rendu compte que […], si dans une publication, quelqu'un mal intentionné prend l'extrait du commentaire et le tape sur [un moteur de recherche…] on […] peut avoir les informations. » (ITW 12).

Complétant la question de l'anonymisation, le débat actuel concernant la publication des données numériques de la recherche met surtout en avant les principes d'ouverture, d'accessibilité, de partage de celles-ci, dans une visée de transparence (Chartron, Schöpfel, 2017). Le mouvement de *l'open science* pousse à la mise en partage généralisée des données et conduit à une injonction à publier

les données de sa recherche, en plus des résultats. Il s'agit de publier des *FAIR data* (*Findable, Accessible, Interoperable, Reusable*), autrement dit des données Faciles à trouver, Accessibles, Interopérables et Réutilisables. « Pour qu'un projet soit financé maintenant au niveau ANR[6] ou au niveau européen, il faut montrer comment le citoyen lambda va pouvoir récupérer nos données, en faire ce qu'il a envie d'en faire, et puis comment d'autres chercheurs peuvent les récupérer aussi. Donc tout ça, ça fait partie de la loi numérique. Il faut rendre accessible les données qui ont été financées sur les… deniers publics. Les jeux de données […] récoltées comme des bases de connaissances, bibliographiques ou cartographiques, ce sont des données qui doivent être mises en accès libre… Il y a énormément de jeux de données sur data.gouv.fr, data.bnf… Tu en as plein » (ITW 9). Pourtant, comme le souligne un interviewé, cette injonction de l'ouverture ne correspond pas forcément à la réalité des SHS. « Un corpus est totalement rattaché à la problématique d'un projet. Il y a aussi peut-être une autre dimension… je pense qu'en sciences sociales, on a intérêt à redire ça parfois parce que l'injonction qu'on a sur l'ouverture des données elle est contraignante, intéressante aussi, mais parfois un peu détachée de ce que c'est finalement la façon dont on construit des objets de recherche » (ITW 7).

L'injonction à fournir des données FAIR laisse entier le questionnement éthique concernant le transfert sociétal des résultats de la recherche et l'ambition de fournir aux enquêtés des données réellement réutilisables. Cette question ancienne relève de la relation enquêteur – enquêté. Elle rappelle que la construction du savoir scientifique se fait avec des catégories d'analyse propres au chercheur qui ne correspondent pas forcément à la vision de l'enquêté. Elle présente néanmoins de nouveaux enjeux et demande un renouvellement de la réflexion. « Dans cette question du transfert sociétal si on veut, maintenant on la traite de plus en plus selon les projets en étant partenaires [insiste sur ce mot] avec les acteurs du monde socio-économique, la société civile… Ce qui fait que ce n'est

6. À noter que « dans le cadre de sa politique science ouverte, l'Agence nationale de la recherche (ANR) demande l'élaboration d'un Plan de Gestion des Données (PGD) pour les projets financés à partir de 2019. Ce document vise à décrire la façon dont les données de recherche seront produites, collectées, réutilisées, stockées, protégées, disséminées ou encore conservées à long terme ». Cette demande de l'ANR suit les recommandations du Comité pour la Science Ouverte (CoSO) qui a attiré « l'attention sur l'adoption d'une démarche progressive qui favoriserait une adhésion des communautés et une meilleure adaptation à l'évolution des pratiques ». https://anr.fr/fr/actualites-de-lanr/details/news/lanr-met-en-place-un-plan-de-gestion-des-donnees-pour-les-projets-finances-des-2019/

pas tant à la fin un transfert, qu'un dialogue qui peut être au fur et à mesure » (ITW 7). « On pourrait fournir des systèmes qui seraient utiles à la société parce qu'ils sont utiles à nous en tant que chercheur. Ça, je trouve que c'est un impensé pour l'instant. [...] c'est-à-dire de mettre à disposition des outils pour faire de la rétro-recherche, comme la rétro-ingénierie » (ITW 3). Pour atteindre cet objectif social, l'ouverture vers des modèles internationaux et plus largement une approche socio-culturelle nous semble être une voie à suivre.

Une dynamique internationale et son ouverture socio-culturelle

La question de l'éthique de la recherche en milieu numérique est en effet traversée par une problématique internationale[7]. D'un côté l'injonction à l'internationalisation des projets de recherche engendre la juxtaposition des contextes scientifiques et réglementaires différents avec lesquels il faut composer : « ça complexifie encore plus, parce que les exigences sont différentes d'un pays à un autre. Et lorsqu'on a des financements, il y a des protocoles très stricts qui sont imposés. Et ces protocoles, eh bien, peuvent changer… Et selon les pays, la mise en place du protocole peut durer très longtemps ! » (ITW 10). D'un autre côté, les différentes réponses et solutions aux problèmes éthiques qui émergent à un niveau international peuvent nourrir une réflexion nationale et/ou disciplinaire.

C'est sur ce modèle que les activités d'associations internationales de chercheurs, notamment celle de *l'Association of Internet Researchers* (AoIR), laquelle jouit d'une grande visibilité dans le champ des *Internet Studies*, ont inspiré la démarche actuelle entreprise au sein de la communauté de la SFSIC en faveur d'un besoin de réflexion avancée concernant l'éthique de la recherche. L'AoIR, par l'intermédiaire de son *ethics working group* – dont le responsable Charles Ess, philosophe de formation, a publié largement sur la question de l'éthique – vient justement de diffuser la troisième version de ses « *guidelines* »[8], à laquelle nous avons contribué. Cette publication, nullement juridique ou normative, est le résultat d'un travail collectif, notamment au sein du groupe de réflexion sur l'éthique, avec une validation finale lors d'un congrès de l'association. Elle est surtout nourrie des réflexions des chercheurs, de leurs expériences de terrain, de leurs

7. Ce point a été abordé dans une publication précédente sur ce sujet, partie refondue (Domenget, Wilhelm, 2018).

8. Cette 3e version a été diffusée lors de son congrès annuel d'automne 2019 (à Brisbane) : https://aoir.org/ethics/.

conceptions multiples du numérique et de son impact. Avec ce travail documenté depuis 17 ans (2002), les chercheurs disposent d'une source de réflexion très utile concernant aussi bien les influences et les évolutions du secteur numérique que le travail des chercheurs s'y intéressant.

En toile de fond de cette publication, se trouve un positionnement qui considère que « l'internet est l'élément médiateur de la vie quotidienne dans les cultures industrialisées et en développement, que nous ayons ou non accès au web. Ainsi, la recherche sur internet doit être considérée dans son sens le plus large » (AoIR, 2012). Cette définition préliminaire ouverte de ce qu'est « la recherche sur Internet » permet d'inclure un spectre très large d'activités de recherche en lien avec le numérique. Selon l'AoIR, la recherche peut être qualifiée comme « recherche sur Internet », quand elle :

> a. utilise Internet pour collecter des données ou des informations, par exemple au moyen d'interviews en ligne, de sondages, de l'archivage ou de moyens automatisés de récupération de données ;
>
> b. étudie la manière dont les gens utilisent et accèdent à Internet, par exemple, en collectant et en observant des activités ou en participant à des sites de réseaux socio-numériques, des serveurs de liste, des sites Web, des blogs, des jeux, des mondes virtuels ou d'autres environnements ou contextes en ligne ;
>
> c. utilise ou participe au traitement, à l'analyse ou au stockage d'ensembles de données, de banques de données et/ou de référentiels disponibles *via* Internet ;
>
> d. étudie les logiciels, le code et les technologies Internet ;
>
> e. examine la conception ou la structure de systèmes, d'interfaces, de pages et d'éléments ;
>
> f. utilise des analyses visuelles et textuelles, des analyses sémiotiques, des analyses de contenu ou d'autres méthodes d'analyse pour observer le Web et/ou les images, les écrits et les formes de média facilités par Internet ;
>
> g. étudie la production, l'utilisation et la régulation d'Internet à grande échelle par les gouvernements, les industries, les entreprises et les forces armées.

Cette typologie très diverse apparaît également dans l'enquête que nous avons menée au sein de la SFSIC. En menant une recherche sur Internet et plus largement sur le numérique, les chercheurs

collectent, traitent une diversité de données et d'informations d'enquête, travaillent également sur un ensemble d'autant plus large de problématiques, telles que les productions et artefacts découverts en ligne (écritures diverses, images, productions amateurs, bases de données professionnelles, etc.), les dispositifs numériques permettant ce travail, leur structuration, les interfaces, etc.

Dans ce même document le comité éthique de l'AoIR a établi une liste de principes qui sont autant de valeurs structurant une démarche éthique dans ce type de recherche, parmi lesquelles la vulnérabilité des sujets qui définit le degré de protection nécessaire et le facteur humain qui est l'*ultima ratio* des recommandations AoIR. En effet, l'être humain est considéré systématiquement au bout de la chaîne de l'action du chercheur, même si celui-ci analyse des objets technologiques ou des données sans co-présence humaine. Dans ce sens, l'AoIR donne une primauté aux droits de l'individu vis-à-vis de la collectivité en cas de doute ou d'éventuels conflits d'intérêts. Ces principes doivent être pris en compte en amont du projet et les décisions doivent être prises de manière collective selon le comité de l'AoIR et après consultation de collègues et de chercheurs. Ces principes, inspirés d'une délibération collective et d'une philosophie humaniste, cherchent actuellement leur égal en France. La SFSIC a voulu s'emparer de ce sujet et combler ce vide.

À partir de cette rapide comparaison avec un exemple international, nous pensons que quelques orientations peuvent guider nos propres démarches. Si les Sic ont l'habitude de prendre en compte la nature évolutive des recherches liées au numérique, il s'agirait notamment de mettre en place une démarche ascendante, organisée autour d'ateliers réguliers, suivant un principe délibératif ; d'ancrer le travail dans une approche interdisciplinaire ; de respecter le cadre situé de chaque problématique éthique et la primauté de l'humain en cas de conflit avec des intérêts collectifs. La question de la *privacy* est un sujet fécond pour les Sic et ses questionnements éthiques. Les liens entre représentations culturelles, usages et discours sociétaux sont particulièrement visibles concernant l'objet de la vie privée, des données personnelles et de la *privacy* (Ess, 2013). La nature transnationale du numérique appelle un travail scientifique à l'échelle internationale pour comprendre les usages et contextualiser les enjeux à ce sujet. À titre d'exemple, si l'IP ne constitue pas pour les États-Unis d'Amérique une donnée personnelle, pour les Européens cette information fait partie des données à protéger. Les législations différentes, comme les réponses des régulateurs, sont à la fois des

sources qui peuvent influencer les positionnements des Sic mais aussi des objets de recherche à part entière. Les différentes représentations et attentes culturelles liées aux données personnelles leur confèrent bien le caractère d'une ressource à titre économique, social et bien évidemment communicationnel. Il s'agit de réfléchir à ce qui peut faire légitimement partie d'une communication, ce qui définit la frontière de l'intimité, de la vie privée et ce qui appartient à l'espace public. Le questionnement éthique n'est alors pas le même quand le périmètre de ce qui est à protéger diffère.

Il en est de même avec les conceptions des Humanités Numériques. Bien qu'il s'agisse d'un mouvement à l'échelle internationale, ses positionnements sont fortement influencés par le cadre socioculturel dans lequel elles évoluent. Cette question est ouvertement posée dans le contexte français dès 2015 (*ibid*., p. 15). Ces auteurs cherchent à éclairer « l'influence réciproque que des traditions nationales de recherche dans ces disciplines peuvent avoir les unes sur les autres » et positionnent le mouvement des Humanités Numériques comme étant en cours de structuration. Ils appellent les sciences humaines et sociales françaises à ne pas rester à l'écart de celui-ci. Ils utilisent la métaphore du chapiteau pour exprimer les liens dans la diversité de ce mouvement et détaillent les stades de développement dans plusieurs pays qui sont reliés à des dynamiques différentes. Bien que présenté comme « un des domaines les plus internationalisés des sciences humaines et sociales » (*ibid.*, p. 23), les auteurs concèdent que le processus d'*évaluation à travers des panels d'experts dans les Humanités Numériques* est fortement anglophone. Exhortant les institutions françaises d'investir plus dans ce « softpower technique », ils avancent que « des problématiques partagées émergent alors, sur les pratiques d'encodage de l'information, sur la structuration, la diffusion et l'archivage des corpus. » (*ibid.*, p. 12). Les Sic ont été depuis longtemps confrontées à ces questions, peu nouvelles pour elles, et il en est de même pour leurs consœurs à l'international.

Conclusion

Si « au plus haut niveau de généralité, on pourrait dire que les Humanités Numériques désignent un dialogue interdisciplinaire sur la dimension numérique des recherches en sciences humaines et sociales, au niveau des outils, des méthodes, des objets d'études et des modes de communication » (Dacos et Mounier, 2015, p. 15) alors chacun de ses aspects implique, comme nous venons de l'exposer, certains choix éthiques. En effet, le choix des outils conditionne

les modalités de vérification de l'impact sur la nature des données. De nouvelles méthodes doivent s'inventer et à leur tour doivent s'évaluer en maintenant un regard critique sur leur portée et leur implication pour les personnes. Les objets d'étude eux-mêmes, bien que (faussement) accessibles, posent de plus en plus de questions à nos enquêtés. Les modes de communication ont eux aussi un impact sur le positionnement éthique comme cela a été exposé plus haut. La *cherchabilité* des médias socionumériques et plus largement du Web expose davantage les données restituées quand celles-ci comportent des données personnelles issues par exemple d'un corpus en ligne.

Une problématique récurrente des recherches en contexte numérique reste le surinvestissement de la portée des observations. L'engouement pour la réalité mesurable par le numérique occulte trop souvent le fait qu'à l'instar de l'échantillonnage numérique lors d'un enregistrement, la « réalité » entière ne peut jamais être observée. Plus encore, la coupure entre l'observation *in situ* et l'observation par le biais des traces numériques est escamotée au profit d'une célébration parfois trop hâtive de la valeur ajoutée des données collectées, dont le contexte de production reste bien souvent sinon opaque, du moins éloigné des besoins de la recherche. La visualisation des données, très présente en Humanités Numériques, produit derrière de beaux visuels des résultats scientifiques plutôt triviaux (Venturini *et al*., 2014) ce qui mène à des justifications théoriques *ad hoc*. Il s'avère donc d'autant plus nécessaire de s'atteler à un nouveau type de travail intermédiaire, qui consiste à comprendre la production des données, à déconstruire les catégories utilisées par les algorithmes et à comprendre comment ils formatent le social à travers ces catégories (Cardon, 2015).

Et si finalement le terme éthique était mal choisi pour parler des sujets qui découlent du numérique ? Les démarches s'intéressant aux questionnements éthiques sont régulièrement critiquées et traitées « de tarte à la crème », de paravent, de procédure administrative lourde, de mot fourre-tout imprécis et moralisateur voire d'antichambre de l'installation de procédures et d'outils de gestion, de contrôle du travail scientifique, d'une perte d'autonomie, etc. Nous pensons que certaines critiques sont à prendre au sérieux et nos retours d'enquête soulignent la crainte exprimée relative au maintien d'une autonomie scientifique. Les *Institutional Review Boards* (IRB) nord-américains sont ainsi souvent pointés du doigt car ils peuvent rapidement devenir un outil administratif pour empêcher des recherches de se faire. Pourtant, face à ce danger, des appréciations plus positives existent également. Patricia Aufderheide (2016) a par exemple publié

un écrit à contre-courant qui souligne les aspects constructifs des échanges issus de ces comités, lesquels permettent de conseiller les chercheurs souvent peu aguerris en matière de principes éthiques. Pour échapper à cet écueil, la mise en place de lieux d'échanges et de discussion collective sur nos pratiques semble être une réponse attendue par les collègues sondés dans l'enquête sur les pratiques de l'éthique en recherche. Nous pensons alors que le travail d'une société savante comme la SFSIC pourrait justement être de donner corps à la notion d'éthique de la recherche, et de faire en sorte qu'existent de tels espaces d'échange constructifs et bienveillants, des lieux collectifs de réflexions et de débats au service des chercheurs. Tel est l'objectif du groupe GENIC présenté en introduction et nous invitons tous les collègues intéressés par ces questions à nous rejoindre.

POSITIONNEMENT

Dans le contexte info-communicationnel numérique actuel, le nécessaire questionnement éthique ne peut se cantonner à la phase de dépôt du projet de recherche et doit avoir lieu pour chaque étape de la recherche :

1. lors de la définition des questions de recherche, interrogeant la responsabilité sociétale des chercheurs et la capacité d'agir des acteurs ;
2. concernant les choix méthodologiques, rendant nécessaire l'adaptation des principes éthiques des méthodes traditionnelles de nature qualitative aux méthodes dites « numérisées » ;
3. en lien avec la production des données de la recherche, leur pertinence et leur traitement ;
4. lors de la phase de diffusion des résultats, en visant le partage de données réellement réutilisables, notamment par les enquêtés.

Cette ambition doit s'appuyer sur la dynamique internationale en cours dans un objectif d'ouverture socio-culturelle afin de s'inspirer des démarches de collectifs confrontés aux mêmes défis. Les réflexions des chercheurs en Sic pourront alors nourrir le questionnement éthique des recherches en Humanités Numériques, dans un objectif d'amélioration de la recherche.

BIBLIOGRAPHIE

Aufderheide P., 2016. *Does This Have to Go Through the IRB?* The Chronicle of Higher Education. Repéré sur le site : https://www.chronicle.com/article/Does-This-Have-to-Go/237476.

AoIR, 2002, 2012, 2019. Ethical decision-making and Internet research: Recommendations from the AoIR ethics working committee. Repéré sur

le site de l'AoiR : https://aoir.org/reports/ethics.pdf ; https://aoir.org/reports/ethics2.pdf ; https://aoir.org/reports/ethics3.pdf.

Balicco L., Broudoux E., Chartron G., Clavier V. et Paillart I. (dir.), 2018. *L'éthique en contexte info-communication numérique*, Louvain-la-neuve, Belgique, De Boeck Supérieur, 160 p.

Boullier D., 2019. *Sociologie du numérique*, Paris, Armand Colin, 184 p.

Boyd D., 2008. *Taken Out of Context. American Teen Sociability in Networked Publics*. Thèse de doctorat en gestion et systèmes d'information, Université de Californie, Berkeley, États-Unis.

Cardon D., 2015. *À quoi rêvent les algorithmes : nos vies à l'heure des* big data, Paris, Seuil, coll. « La République des idées », 112 p.

Chartron G. et Schöpfel J., 2017. « *Open access* et *Open science* en débat ». *RFSic*, n° 11. https://journals.openedition.org/rfsic/3331. Dernière consultation le 12/03/2020.

Dacos M. et Mounier P. (2015). *Humanités Numériques*. Hal. https://hal.archives-ouvertes.fr/hal-01228945. Dernière consultation le 12/03/2020.

Domenget J.-C. et Wilhelm C., 2018. « L'éthique des pratiques de recherche liées au numérique en Sic : le rôle de la société française des sciences de l'information et de la communication » dans L. Balicco *et al.* (dir.). *L'éthique en contexte info-communicationnel numérique*. Louvain-la-neuve, Belgique, De Boeck Supérieur, pp. 101-111.

Ess C., 2013. *Digital Media Ethics*, Londres, Angleterre, Polity Press, 200 p.

Ess C., 2002. « Introduction. Special Issue on Internet Research Ethics ». *Ethics and Information Technology*, 3 (4), pp. 177-188

Iphofen R. (ed.), 2017. *Advances in research ethics and integrity: Finding common ground – Consensus in research ethics across the social science (vol. 1)*, UK, Emerald group publishing limited, 304 p.

Latzko-Toth G. et Pastinelli M., 2013. « Par-delà la dichotomie public/privé : la mise en visibilité des pratiques numériques et ses enjeux éthiques ». *Tic&Société*, 2 (7). DOI : https://doi.org/10.4000/ticetsociete.1591. Dernière consultation le 12/03/2020.

Mounier P., 2018. *Les Humanités Numériques*, Paris, Éditions de la Maison des sciences de l'homme. https://books.openedition.org/editionsmsh/12006. Dernière consultation le 12/03/2020.

Simonnot B., 2018. « Conduire des recherches en régime numérique : vers un cadre conceptuel de réflexion éthique » dans L. Balicco *et al*. (dir.). *L'éthique en contexte info-communicationnel numérique*. Louvain-la-neuve, Belgique : De Boeck Supérieur, pp. 11-21.

Szoniecky S. et Safin S., 2017. « Modélisation éthique de l'Internet des Objets ». *Internet des objets*, 2 (1). https://www.openscience.fr/Modelisation-ethique-de-l-Internet-des-Objets. Dernière consultation le 12/03/2020.

Théviot A., 2013. « Devenir "ami" avec 4 500 enquêtés. Les enjeux éthiques de l'analyse d'interfaces semi-privées ». *Tic&société*, 2 (7). DOI : 10.4000/ticetsociete.1608. Dernière consultation le 12/03/2020.

Venturini T., Cardon D. et Cointet J.-P., 2014. « Présentation ». *Réseaux*, n° 188 (6), pp. 1-20. https://doi.org/10.3917/res.188.0009. Dernière consultation le 12/03/2020.

Vitali-Rosati M., 2012. Une éthique appliquée ? *Éthique publique* 2 (14). http://journals.openedition.org/ethiquepublique/995. Dernière consultation le 12/03/2020.

Wieviorka, M., 2013. *L'impératif numérique ou la nouvelle ère des sciences humaines et sociales ?*, Paris : CNRS Éditions , 60 p.

Woodfield K. (ed.), 2017. *Advances in research ethics and integrity: The Ethics of Online Research (vol. 2)*, UK, Emerald group publishing limited, 272 p.

CHAPITRE VI

Inscrire les Humanités Numériques dans une économie de la production scientifique

Vincent Bullich
Université Grenoble Alpes, Gresec EA 608

Résumé : Par la mobilisation d'outils, de méthodes et terrains, d'approches et de concepts inédits, les Humanités Numériques posent à nouveaux frais la question de l'interprétation et, plus largement, celle des conditions de possibilités du sens. Ce faisant, elles soulèvent un ensemble d'enjeux à la confluence de l'épistémologie et du politique. Le principal se rapporte aux processus mêmes de production et d'analyse des données de la recherche. La démarche de quantification, désormais assistée informatiquement, puis de configuration des énoncés scientifiques dont elles relèvent appellent un examen des procédures de mise en forme et de traitements de matériaux étudiés. Du point de vue des sciences de l'information et de la communication, cet examen passe par l'étude des « médiations organisées », c'est-à-dire par la prise en compte de l'épaisseur propre aux dispositifs sémio-techniques et par celle des stratégies des acteurs industriels ou institutionnels qui les mettent en œuvre.

Mots-clés : Pratiques scientifiques, herméneutique numérique, industrialisation des dispositifs, marchandisation de la recherche.

Abstract: By mobilizing new tools, methods and fields, approaches and concepts, digital humanities are once again raising the question of interpretation and, more broadly, that of the conditions of possibility of meaning. In doing so, they raise a range of issues at the confluence of epistemology and politics. The main one relates to the very processes of producing and analyzing research data. The quantification process, now computer-aided, then the configuration of the scientific statements to which they relate calls for an examination of the procedures for shaping and processing the materials studied. From the point of view of information and communication Sciences, this examination passes by the study of "organized mediations", that is to say by taking into account the thickness proper to "semiotechnical" devices and by that of the strategies of the industrial or institutional actors who implement them.

Keywords: Scientific practices, digital hermeneutics, industrialisation of scientific devices, commodification of research.

L'histoire des Humanités Numériques est indissociable de celle des innovations industrielles dans le domaine de l'informatique. Les historiens ou experts du domaine s'accordent ainsi pour situer l'origine de cette approche, quelquefois qualifiée de « transdiscipline »[1], dans la démarche du jésuite italien Roberto Busa qui scella une alliance inédite en sciences humaines et sociales entre un établissement universitaire, en l'occurrence le Massachussetts Institute of Technology, et une entreprise d'informatique, IBM (Berra, 2012). D'emblée, le projet scientifique, qui sera un demi-siècle plus tard désigné par l'expression de « *Digital Humanities* », est le produit de la collaboration entre le monde de la recherche académique et celui de la recherche et développement industriel. Par conséquent, une caractéristique essentielle de l'approche est la puissance heuristique conférée à « l'outillage » numérique. Le caractère décisif des instruments dans la production de connaissances n'est pas en soi nouveau, la science ayant toujours été tributaire d'un « ordre matériel » qui configure les conditions de savoir (Waquet, 2015). La nouveauté (relative) vient plutôt de la matérialité singulière du « numérique ». Tel est le point de départ de la réflexion que ce court texte expose. Celui-ci a un double objet. Il s'agit dans un premier temps de considérer les effets de l'intégration systématique de dispositifs numériques aux procédures de recherche et plus particulièrement à celles concernant l'interprétation (qui est au cœur des Humanités). Dans un second temps, le propos s'élargit aux conditions de production et de mise en œuvre de ces dispositifs, abordées par le prisme des processus d'industrialisation et de marchandisation.

Penser les médiations et la matérialité numérique dans leur multi-dimensionnalité

L'attention portée aux médiations est au cœur des approches en Sic depuis la création de la discipline. Appréhendée comme une activité ou une action sur le cours même d'un processus communicationnel,

1. Manifeste des Humanités Numériques 2010, reproduit dans : Mounier, 2010.

la médiation se manifeste au travers de traits multiples qu'il appartient au chercheur d'identifier. Ces traits sont généralement rapportés, dans un premier temps, aux moyens employés. Par exemple, dans le dossier consacré aux Humanités Numériques de la *Revue Française des Sciences de l'information et de la communication*, Jean-Claude Domenget, Julia Bonaccorsi et Valérie Carayol (2016) rappellent que « la question de l'artefactualité technique, impliquée par toute transmission ou transformation d'informations ou de connaissances, est depuis la fin des années 70 une question essentielle des Sic » (p. 1). L'analyse vise par conséquent à rendre compte avant toute chose de l'action inhérente aux dispositifs mobilisés dans les pratiques scientifiques.

De la même façon, la plupart des hérauts des Humanités Numériques opèrent un retour réflexif sur leurs pratiques en se centrant, précisément, sur la médiation technique. Mieux, certains partagent une approche résolument « matérialiste » et considèrent ainsi que « c'est […] la contrainte imposée par le traitement numérique de l'information au chercheur qui le conduit à formaliser le matériau sur lequel il travaille et le raisonnement qu'il mobilise » (Mounier, 2018, p. 60). Suivant cette perspective, les outils et machines numériques ne constituent pas une simple instrumentation de l'activité scientifique mais déterminent bien plutôt une heuristique nouvelle : les Humanités Numériques désigneraient en cela un mode singulier d'appréhension des phénomènes – procédant d'une intrication entre mesures et représentations, calculs et enjeux de sens, computation et interprétation – à même de renouveler tout un pan de la recherche en SHS (Dacos, Mounier, 2014).

Ce positionnement ne saurait manquer d'interpeller les chercheurs en Sic. En effet, l'instanciation matérielle de l'information, puis l'instauration de cette information en données scientifiques sont, depuis des années, des objets d'études info-communicationnelles. De même, les travaux sur la numérisation documentaire et les systèmes d'informations informatisés n'ont (heureusement) pas attendu la vague des Humanités Numériques pour se développer. En outre, nombre d'études en Sic se sont construites contre l'invisibilisation des médiations qui semble régulièrement être l'objectif de ceux qui les mettent en place. De nombreux points de jonction sont ainsi aisément identifiables, et, hormis certains cas-limites[2], on aurait tort d'opposer *a prio-*

2. Certaines approches s'accordent en effet mal avec l'épistémologie des Sic. On peut, par exemple, citer le projet scientifique « Culturomics » (Aiden, Michel, 2014) qui consiste en une « lexicologie computationnelle », c'est-à-dire une approche quantitative de l'étude de la culture humaine au travers d'une analyse informatisée de très grands corpus de

ri les approches des Humanités Numériques à celles qu'on trouve en Sic.

Néanmoins, ces dernières proposent une perspective nettement plus étendue en ce qu'elles considèrent non seulement les dispositifs participant de la recherche scientifique mais également leurs conditions de production. La focale se resserre donc et le champ s'élargit : le regard englobe dès lors un ensemble de phénomènes laissés de côté, semble-t-il, par les Humanités Numériques. Cet élargissement du périmètre des études permet d'appréhender la configuration matérielle des pratiques ainsi que les médiations produites en les rapportant aux « dimensions » multiples de l'activité, c'est-à-dire à la pluralité des ordres de détermination et des plans d'analyse. Celles-ci se conjuguent, s'articulent de façon plus ou moins harmonieuse ou plus ou moins conflictuelle, et bornent, par le jeu de leurs interactions, le contexte de production. En cela, l'approche adoptée dépasse l'examen épistémologique *stricto sensu*, et l'exigence de multi-dimensionnalité rapporte sur le plan économique, politique et social, les conditions et produits de la recherche. Afin d'illustrer la portée de la démarche, nous allons nous intéresser dans les lignes suivantes à la phase cruciale dans l'élaboration des connaissances qu'est l'interprétation.

La délégation logicielle de l'interprétation

Les Humanités Numériques se caractérisent par le rôle central conféré aux logiciels et agents logiciels[3] dans leurs méthodologies. Cette « logicialisation »[4] est observable tant dans les procédures d'obtention et de traitement des données, que dans celles de production puis restitution des résultats. Du point de vue de l'étude des médiations, tout l'enjeu réside alors dans la capacité d'action et le degré d'autonomie de ces agents. Se présente en effet une situation inédite à partir du moment où certaines tâches qui participent de l'élaboration

textes généralement répartis sur une longue période. Le pari qui est fait est que le volume de données traitées serait à même de révéler des tendances imperceptibles par des méthodes historiques « classiques » (suivant le principe des « signaux faibles »). Or, la décontextualisation radicale que les chercheurs qui proposent cette approche font subir aux unités lexicales qu'ils manipulent laisse perplexe quant à leurs présupposés théoriques et à la portée de leur projet : ressuscitant grâce aux instruments numériques la démarche structuraliste, ceux-ci visent, ni plus ni moins, à produire une « science prédictive de l'histoire » (notre traduction de Aiden, Michel, 2014, p. 211). L'oxymore est savoureux.

3. Programmes qui agissent de façon autonome pour l'accomplissement d'une tâche.

4. Selon la formule de l'économiste François Horn (2000) qui désigne ainsi la part continûment croissante des logiciels dans les activités de production de biens et services et celles liées à leur valorisation.

interprétative sont dévolues à un dispositif à même de (ré)agir de façon autonome aux commandes ou requêtes qui lui sont adressées (en fonction d'un paramétrage initial, complété quelquefois d'une disposition d'auto-apprentissage). Cette exécution automatique – caractéristique du passage de l'outil à la machine (Ingold, 2010) – se traduit évidemment par une complexité et une épaisseur accrues des médiations en jeu. C'est cette affirmation de la décision machinique dans le processus interprétatif qui est désignée ici par le syntagme de « délégation logicielle »[5].

Cette herméneutique[6] « assistée » ou « conduite »[7] informatiquement soulève de nombreuses interrogations. D'emblée, le projet même d'une computation généralisée a été questionné. Les Humanités Numériques visent en effet à la compréhension des phénomènes par substitution de variables discrètes aux valeurs continues, puis par substitution du calcul probabiliste à une poïétique du sens en contexte. Face à cette démarche, de nombreux chercheurs ont procédé à une mise en garde contre l'excès de confiance dans la force objectivante des données ou contre l'hégémonie scientifique de la « conception quantitative de la preuve » (Doueihi, 2015, p. 705). Rapporter l'interprétation à un produit de la statistique bayésienne ne va en effet pas sans une déperdition importante des qualités du singulier. La réduction de tout phénomène aux mesures chiffrées, cette mise en compatibilité pour une mise en comptabilité permet certes de s'abstraire en partie de l'équivocité du langage ; elle présente également le risque de ramener toute propriété au commensurable et ainsi d'exclure de l'analyse les manifestations idiosyncratiques[8].

Quoi qu'il en soit, l'autorité du formalisme « logico-mathématico-informatique » est censée répondre à l'imperfection langagière ; l'accès facilité techniquement à des volumes inédits de données ainsi que leur analyse algorithmique sont censés suppléer, quant à eux, les limites de l'interprétation « classique ». La conjonction des deux facteurs a favorisé l'apparition de discours provocants – et pour tout dire saugrenus – sur la « fin de théorie » et son remplacement par

5. Nous nous référons ici au concept de « délégation technique » proposé par Bernhard Rieder (2006) qui caractérise ainsi une relation homme-machine qui n'est pas que de l'ordre de la « pure » substitution mécanique (c'est-à-dire, neutre axiologiquement) mais qui, au contraire, a des incidences « politiques, culturelles et sociales » sur les activités humaines (Rieder, 2006, p. 113).

6. Dans une acception large : comme procédure méthodologique non pas exclusivement cantonnée à l'étude des textes mais intégrant toute démarche scientifique.

7. Suivant le degré d'intervention des agents logiciels.

8. Cet écueil est accentué par le caractère automatisé des dispositifs mis en place, voir *infra*.

la détection automatisée des « signaux faibles » dans « le déluge » annoncé des données (Anderson, 2008)[9], ou sur les possibilités, au travers de leur traitement statistique, d'identification de « lois déterministes » de la culture (Aiden, Michel, 2014). Cette aspiration à automatiser l'épiphanie et à faire des Humanités des sciences nomothétiques est tout à fait symptomatique de l'affirmation du technocentrisme qui menace les Humanités Numériques, mais qu'il est aisé de neutraliser épistémologiquement.

Toutefois, la délégation logicielle a une incidence plus insidieuse sur l'interprétation. En effet, l'abstraction est inhérente à tout processus interprétatif, mais celle-ci est d'une nature particulière à partir du moment où elle est gouvernée par l'ordre informatique et la « raison computationnelle » (Bachimont, 2010). Elle ne consiste dès lors plus uniquement en une « amputation » de la réalité suivant une logique synthétique mais révèle au contraire des éléments sémiotiques nouveaux, produits par le dispositif sans que le chercheur n'en soit (forcément) conscient. En effet, outre l'écueil des points aveugles liés à l'évacuation des idiosyncrasies, la production automatisée du commensurable a pour effet systématique l'adjonction incontrôlée d'éléments à interpréter ou d'éléments qui apparaissent comme « déjà interprétés ». Cela est notamment dû au fait que les dispositifs impliqués ont tendance moins à détecter puis représenter les qualités des objets analysés qu'à les définir eux-mêmes, c'est-à-dire moins à discerner des propriétés qu'à les déterminer selon une loi normale. Incapables de « se taire », ils fonctionnent en cela comme des instances de qualification, attribuant par eux-mêmes des caractéristiques aux objets de l'analyse. Cette tendance est renforcée par les procédés d'apprentissage automatique qui conduisent, idéalement, à l'autodétermination des règles à appliquer lors du traitement de ces objets. C'est donc le « point de vue » des dispositifs, déterminé à partir des éléments antérieurement traités, qui conditionne l'objet et la nature du traitement subséquent. Cette logique auto-renforçante est à l'évidence susceptible de produire un biais interprétatif majeur.

Les manifestations numériques d'un objet d'étude sont ainsi particulièrement labiles et subissent des transformations multiples : au travers de la mise en données, au travers de l'intégration de celles-ci à des corpus/bases, au travers des opérations algorithmiques de traitement, puis au cours du passage de fichiers plats en représentations

9. Indiquons ici que Chris Anderson ne se réfère pas dans son article aux Humanités Numériques mais à toutes les démarches de recherche.

visuelles structurées (depuis la « simple » interface aux datavisualisations les plus sophistiquées). Chacune de ces transformations est le produit d'un logiciel qui agit comme instance à la fois qualificative et interprétative. Le biais potentiel induit par l'automatisation de ces procédures concerne par conséquent tout le processus herméneutique : de la sélection et l'obtention des données – qui ne sont jamais « brutes » mais toujours « *in*-formées » dans un but préalable et donc porteuse d'une intentionnalité et d'éléments contextuels (Gitelman et Jackson, 2013) – jusqu'à la restitution des résultats (*a fortiori* quand celle-ci opère par déplacement du plan de l'écrit à celui du graphe).

Ces aspects ont déjà pour une bonne part été étudiés et les médiations informatiques de l'interprétation sont généralement au cœur des travaux concernant les Humanités Numériques[10]. L'objectif est dès lors de prolonger cet effort de réflexivité, en dépassant le cadre de son application aux pratiques des chercheurs configurées par les dispositifs mobilisés et en s'intéressant désormais aux conditions de production de ces mêmes dispositifs.

L'industrialisation de la logistique scientifique

Les médiations qu'appréhende la recherche en Sic sont « organisées » (Miège, 2004). Cela veut dire qu'elles sont non seulement déterminées par un ordre matériel et une rationalité afférente, mais qu'elles sont également et nécessairement la concrétisation d'intentionnalités, d'intérêts et de capacités d'agir, qu'elles résultent toujours d'actions dirigées vers un but. En cela, les médiations sont, ainsi que le souligne Yves Jeanneret (2014), porteuses d'un « enjeu de pouvoir ». Si la technique est centrale, l'exigence de multi-dimensionnalité nous amène à prendre en compte d'autres ordres de détermination, d'autres rationalités à même de rendre compte de la matérialité des dispositifs ainsi que des pratiques observées. Les Sic présentent en cela une position épistémologique forte que l'on pourrait formuler de façon aphoristique par « pas de technique ni de signe en dehors du social ». L'imprégnation sociale de la technique étant polymorphe, nous allons ici nous concentrer sur deux aspects : les déterminants économiques et leurs conséquences politiques.

La « logicialisation » de la recherche en Humanités est la partie émergée de « l'iceberg » technique. Elle concentre généralement l'attention

10. Nous renvoyons le lecteur au dossier déjà cité de la *Revue française des sciences de l'information et de la communication* (2016/8), ainsi qu'à ceux des revues *Questions de communication* (2017/31), *Critique* (2015/819-820) et *Variations* (2016/19).

en ce qu'elle est au plus près des pratiques. Ces logiciels s'inscrivent toutefois dans un système technique plus vaste, le dispositif évoqué jusque-là étant dépendant d'infrastructures télécommunicationnelles et énergétiques. On peut ainsi considérer une logistique scientifique nouvelle qui voit le jour dans le domaine des Humanités sous l'impulsion de leur « numérisation ». Composée notamment d'un ensemble de programmes et de terminaux, de réseaux et de modalités de connexions, de racks de stockage (centres de données) et de commutateurs, cette logistique configure codes et signes et par extension les pratiques de recherche et leur communication. Elle manifeste le tournant proprement industriel de l'heuristique des Humanités.

Celui-ci est en partie le résultat de stratégies d'acteurs économiques. On peut en distinguer trois catégories : les opérateurs de télécommunication, les fournisseurs de solutions logicielles et les producteurs de matériel. Interdépendants du point de vue de la production, ces acteurs co-opèrent également à des degrés divers à la production scientifique. Ce n'est donc évidemment pas un hasard si, depuis 2016, se tient chaque année le *Digital Humanities Infrastructure Symposium* à l'Université de Californie à Los Angeles (UCLA). Préalablement, le *Manifeste des Digital Humanities*[11] appelait, par exemple, à la construction de « cyberinfrastructures[12] évolutives répondant à des besoins réels » (§ 14). Cette question logistique est d'autant plus sensible que des industriels de premier plan ont profité de l'engouement récent des chercheurs en sciences humaines et sociales pour l'instrumentation numérique pour se positionner dans le monde académique. La situation n'est en rien inédite, mais s'est spectaculairement amplifiée depuis l'affirmation des Humanités Numériques comme programme « transdisciplinaire »[13].

11. Ce manifeste a été rédigé en 2010 à Paris et cosigné par plus de 250 chercheurs de 20 laboratoires internationaux de la conférence « ThatCamp » (The Humanities and Technology Camp).

12. Ainsi que l'expose Pierre Mounier (2010, § 1), cette notion apparait initialement dans un rapport du Council of American Learned Societies afin de « désigner un ensemble d'informations, d'expertises, de standards, de stratégies, d'outils et de services qui sont partagés largement entre les communautés mais développés spécifiquement pour des usages savants ». En cela, la « cyberinfrastructure » constitue une circonscription d'éléments plus réduite que l'ensemble dispositifs et infrastructure mentionné *supra*. Elle se rapporte néanmoins foncièrement à l'idée de logistique dont il est ici question.

13. Ainsi que le souligne Dominique Vinck (2016), cette affirmation s'est réalisée en plusieurs étapes, au travers notamment de la publication de chartes, déclarations et manifestes portés par des sociétés savantes ou associations internationales (dont le Manifeste évoqué supra et infra constituent le parangon). Ces publications ont visé à la fois la promotion au sein du monde académique d'une approche nouvelle, qui cherche à fédérer

IBM puis Hewlett Packard ont, par exemple, de longue date investi les universités, pas uniquement les départements informatiques mais également ceux consacrés aux SHS. Si l'assistance informatique s'est développée dès la décennie 1980, notamment pour les disciplines s'appuyant sur des méthodes quantitatives, il s'agit de souligner ici le changement d'échelle induit par les approches des Humanités Numériques. Celui-ci est perceptible bien évidemment dans les volumes de données traités et dans les capacités computationnelles de traitement. Celui-ci est perceptible également dans l'extension et la superposition des agencements matériels et logiciels. Il en résulte bien souvent un étirement de la chaîne de co-opération de la production scientifique, au sein de laquelle interviennent de plus en plus d'acteurs industriels. La situation américaine est à cet égard édifiante. Les entreprises y ont en effet multiplié leurs implications dans des projets estampillés « Humanités Numériques » : opérateurs de télécommunication (comme AT&T[14] ou Verizon[15]), fabricants de matériel (comme Hewlett Packard[16]), de solutions logicielles (comme IBM[17]) et, évidemment, les géants de l'informatique et du Web (Microsoft[18] et Google[19] en tête). La plupart des acteurs dominant l'économie numérique sont donc présents sur ce volet nouveau de la recherche universitaire que sont les *Digital Humanities*. Leur implication est en

et à s'insérer dans des cadres institutionnels ; ils ont visé également à la reconnaissance de la démarche par les acteurs économiques à même de contribuer au financement et à la valorisation de la recherche : commanditaires et bailleurs de fonds publics comme privés, prestataires techniques, partenaires commerciaux.

14. Dans le cadre, par exemple, des activités de l'AT&T Learning Studio de l'Université d'Abilene.

15. Dans le cadre, par exemple, d'un partenariat pour la mise en place des ateliers « Digital Humanities » du National Endowment for the Humanities.

16. L'entreprise a, dès le début des années 1970, contribué au financement et à la réalisation du Thesaurus Linguae Graecae, la collection complète des textes en grec ancien. Désormais, Hewlett-Packard contribue financièrement et techniquement à la mise en place de programmes d'enseignement « numérique » des Humanités, en lien avec un groupe de chercheurs issus des plus prestigieuses Universités américaines (Harvard, le MIT, Yale, etc.).

17. IBM propose différentes bourses ou prix aux chercheurs en Humanités Numériques. En outre, la société participe au financement de conférences internationales et, ce, depuis la conférence « Computers for the Humanities » de 1965 qui sera à l'origine du Yale Digital Humanities Laboratory.

18. Microsoft développe ainsi par exemple un programme nommé Digital Heritage and Humanities conduit en interne par les laboratoires de la société et en lien avec des équipes universitaires (voir, entre autres, le Brown University's Garibaldi Project : https://library.brown.edu/cds/garibaldi/).

19. Google propose notamment plusieurs bourses de recherche dans le cadre des « Google Digital Humanities Research Awards ». En outre, la société offre depuis 2010 des financements plus conséquents pour des projets longs (voir *infra*).

outre appelée à se renforcer en raison d'un recours de plus en plus fréquent au *Cloud Computing* : en multipliant les applications, serveurs et connexions-réseaux, ce type de procédés renforce l'intervention des leaders mondiaux du domaine que sont Amazon, Google et Microsoft (Mosco, 2016). Enfin, cette logistique nouvelle, éminemment industrielle, favorise non seulement un changement d'échelle mais également, et de façon croissante, un changement de nature dans le matériau même des études. Il est devenu ainsi courant que ce matériau soit lui-même un produit industriel (plus précisément : des données pré-traitées et produites industriellement). Par exemple, le programme Culturomics est né d'une alliance entre de jeunes chercheurs de Harvard, du MIT et de Princeton et Google, la démarche s'appuyant sur le titanesque travail de numérisation de livres commandité par la société de Menlo Park. Il en va de même pour l'ensemble des études réalisées à partir de *tweets* ou celles sur des données extraites des interfaces de programmation. Cette entremise croissante des industriels du numérique n'est pas sans incidence sur les exigences scientifiques. Outre les écueils liés aux limites techniques (un défaut de capacité de calcul, des lacunes dans les données disponibles ou dans leur compatibilité, etc.) ou alors imputables aux chercheurs mêmes (dans la définition des objectifs de recherche, la méthodologie idoine ou sa traduction informatique, etc.), les médiations propres au caractère industriel de la logistique mobilisée devient un enjeu épistémologique ; *a fortiori* parce que cette industrialisation de la logistique scientifique (mais également du matériau étudié) s'accompagne inéluctablement d'une intensification de la pression marchande sur les instruments et les produits de la recherche.

Les risques d'une marchandisation accrue de la recherche

L'industrialisation traduit une mutation de la production généralement motivée, dans les économies capitalistes, par une recherche de profit. En cela, le projet de marchandisation précède la mise en place d'une organisation industrielle ; en retour, cette forme d'organisation productive accélère le devenir marchand des activités concernées.

La recherche n'est évidemment pas épargnée par cette tendance de fond. La finalité marchande, cette volonté de conversion de toute chose en valeur d'échange, a investi depuis longtemps les cénacles universitaires, à commencer, une nouvelle fois, par les établissements américains. Depuis la promulgation du Bayh-Dole Act en 1980, les universités d'outre-Atlantique peuvent en effet déposer et exploiter des brevets, même si la recherche a été financée par des fonds fédéraux

(et donc publics). Cette tendance à la production et à la valorisation des ressources propres issues de l'activité scientifique s'est prolongée au cours des décennies suivantes. La recherche scientifique s'est alors scindée en recherche destinée à produire un savoir « commun » d'une part, alors labellisée comme « fondamentale » et principalement soutenue par des fonds publics, et, d'autre part, une recherche ayant une finalité explicitement économique, « appliquée » et reposant sur des financements mixtes publics/privés (Peters, Marginson, Murphy, 2009).

Cette extension du domaine marchand de la science n'affecte évidemment pas uniquement les Humanités Numériques. Cependant, celles-ci sont particulièrement enclines à donner prise au processus. Ceci, premièrement, parce que les méthodes employées, reposant largement sur une philosophie du « faire » concrétisée dans des adjuvants techniques (Le Deuff, Clavert, 2014), sont à même de générer des brevets exploitables commercialement (ces dispositifs étant potentiellement transposables à d'autres domaines). En cela, les liens entre monde académique et secteurs industriels apparaissent renforcés et la frontière séparant les activités de plus en plus poreuse : de nombreux exemples aux États-Unis illustrent la façon dont les universités deviennent l'antichambre de l'industrie, les laboratoires fonctionnant comme incubateurs de *start-up*, y compris dans les domaines SHS[20].

Deuxièmement, les données produites sont à même de faire l'objet d'une appropriation marchande. L'appétence des firmes dominantes de la Silicon Valley pour celles-ci ne saurait décliner et la recherche pourrait agir à cet égard comme un auxiliaire involontaire, à plus forte raison si les chercheurs utilisent des dispositifs dont elles sont les propriétaires. À l'instar de ce que l'on observe pour les plateformes numériques grand public (réseaux socio-numériques, plateformes de diffusion audiovisuelle ou vidéoludique, etc.), cette captation se réaliserait ainsi par le truchement de dispositions contractuelles généralement négligées par les utilisateurs. Évidemment, l'appropriation des données est un enjeu qui dépasse largement le cadre scientifique et il semblerait que les pouvoirs publics commencent à prendre la mesure du risque (ainsi qu'en témoigne par exemple le Règlement général sur la protection des données en Europe). Toutefois, les précédents liés à

20. C'est déjà fréquemment le cas : par exemple, l'Infolab, laboratoire d'informatique et de sciences du langage, de l'Université Northwestern dans l'Illinois a été l'incubateur de plusieurs sociétés, dont Narrative Science qui propose des solutions logicielles d'écriture automatique à destination des entreprises de presse notamment, se présentant ainsi comme un précurseur du « robot-journalism ».

la brevetabilité du vivant et la privatisation des connaissances génétiques ne font que légitimer ce type de craintes.

Un troisième enjeu réside dans les modalités de publicisation scientifique. Les promoteurs des Humanités Numériques sont sur ce sujet sans équivoque. Ainsi que le stipule au travers d'un jeu de mot le *Manifeste des Humanités Numériques 2.0*[21] de 2009 : « Intellectual property must open up, not close down the intellect and proprius » (§ 18). Suivant ce précepte, il s'agit non seulement de réinventer les pratiques d'écritures scientifiques, sclérosées par des normes d'auctorialité et de publications jugées obsolètes, mais également de « libérer » le fonctionnement de la propriété intellectuelle « de l'emprise du Capital »[22]. S'énonce ici clairement une tactique de résistance à la marchandisation au travers des programmes liés à la « science ouverte » et aux « communs de la connaissance ». Si l'intention est louable et la position franche, les effets de celles-ci apparaissent, eux, plus ambivalents. À l'heure actuelle, l'intensification du processus marchand s'observe à la fois dans les mutations de l'édition scientifique (au travers notamment de la reconfiguration des chaînes de valeur et de la structure désormais nettement oligopolistique du marché mondial) et dans des pratiques nouvelles et résolument commerciales de diffusion des savoirs (au travers notamment de l'utilisation de plateformes audiovisuelles généralistes ou spécialisées dans les MOOC). Si la domination souveraine de quelques éditeurs sur la communication scientifique est largement dénoncée au sein de la communauté académique, les idéaux d'une science débarrassée de la mainmise de ceux-ci se heurtent, pour leur part, à un autre mur de réalité économique. En effet, la mise en œuvre du projet d'« ouverture » des sciences se traduit pour le moment par un double mouvement associant « logiques de fragmentation et montée des intermédiaires » (Dillaerts, Boukacem-Zeghmouri, 2018). Par conséquent, l'opposition aux logiques « propriétaires » et les objectifs de libre diffusion des savoirs scientifiques conduisent non pas forcément à une évacuation des finalités économiques gouvernant la publicisation scientifique mais bien plutôt à une reconfiguration du jeu concurrentiel et un déplacement de son centre de gravité. La portée d'un tel phénomène est synthétisée dans la question cinglante que posent Joëlle Farchy et

21. The Digital Humanities Manifesto 2.0. Récupéré du site : http://manifesto.humanities.ucla.edu/2009/05/29/the-digital-humanities-manifesto-20/

22. « [...] copyright and IP standards must be freed from the stranglehold of Capital », The Digital Humanities Manifesto 2.0, § 13.

Pascal Froissart (2010, p. 143) : « Est-il plus efficace pour diffuser la recherche d'assurer une rente de monopole à Google ou à Elsevier ? ».

Contre cette tendance, les initiatives de la part de chercheurs eux-mêmes se sont multipliées, souvent soutenues par des institutions publiques : on peut citer, pour la France, les actions du Centre pour l'édition électronique ouverte, dirigé par Marin Dacos et Pierre Mounier et au niveau Européen, les *consortia Digital Research Infrastructure for the Arts and Humanities* (DARIAH) et *European Research Infrastructure for Language Resources and Technology* (CLARIN ERIC). En outre, de plus en plus de chercheurs sont attentifs au devenir des données numériques et de nombreuses chartes éthiques ont ainsi été élaborées au sein d'établissements, de réseaux de recherche nationaux et internationaux afin, entre autres, de prévenir les appropriations indues. Ces exemples illustrent un geste politique de la part d'acteurs de la recherche, individuels et collectifs ou institutionnels, qui élaborent des programmes en opposition à cette tendance de fond qu'est la marchandisation. Il s'agit donc de prolonger cet effort de vigilance visant à ce que la numérisation des Humanités ne soit pas le moment d'une immixtion d'intérêts extra-scientifiques dans le cours des sciences, que les dispositifs mis en place ne fonctionnent pas comme le cheval de Troie des industries du logiciel notamment. Une nouvelle fois, l'enjeu est tout autant scientifique que politique : il s'agit de définir un modèle de développement de la recherche qui ne soit pas dépendant d'une logistique produite dans une finalité exclusivement lucrative afin de garantir son autonomie. Pour cela, l'attention portée à l'action des dispositifs numériques nécessite de se doubler d'une perspective en termes d'économie politique de la production scientifique. Bref, elle implique une démarche, à la fois multi-dimensionnelle et interdisciplinaire, telle que privilégiée par les Sic, qui apparaissent, pour cette raison, à même d'agir comme sentinelle.

POSITIONNEMENT

Les approches regroupées sous l'appellation « Humanités Numériques » se caractérisent par l'adjonction systématique de dispositifs sémio-techniques aux méthodes de recherche et d'interprétation des résultats en SHS. Indispensables à la démarche heuristique, ces dispositifs manifestent également un processus d'industrialisation de la logistique scientifique appliquée aux Humanités. Celles-ci deviennent dès lors partiellement dépendantes des acteurs économiques qui les conçoivent et les mettent en œuvre, et dont les intérêts ne convergent pas forcément avec ceux des scientifiques. Pour cette raison, « questionner les Humanités Numériques » implique de questionner les conditions de production des dispositifs mobilisés, tout autant que celles des savoirs et savoir-faire. L'examen épistémologique répond dès lors à une exigence de multi-dimensionnalité : il s'agit d'appréhender dans un même mouvement théories et méthodes, dispositifs et pratiques, discours et médiations et d'inscrire ceux-ci dans une économie générale de la production scientifique.

BIBLIOGRAPHIE

Aiden E. et J.-B. Michel, 2014. *Uncharted: Big Data as a Lens on Human Culture*, New York, Riverhead Books, 288 p.

Anderson C., 2008. « The End of Theory: The Data Deluge Makes the Scientific Method Obsolete ». *Wired*. https://www.wired.com/2008/06/pb-theory/. Dernière consultation le 12/03/2020.

Bachimont B., 2010. *Le sens de la technique : le numérique et le calcul*, Paris, Les belles Lettres, 192 p.

Berra A., 2012. « Faire des Humanités Numériques » dans P. Mounier (éd.). *Read/Write Book 2*. Marseille, OpenEdition Press, pp. 25-43.

Dacos M. et P. Mounier, 2014. *Humanités Numériques.* État des lieux et positionnement de la recherche française dans le contexte international. Paris, Institut Français/OpenEdition. https://www.enssib.fr/bibliotheque-numerique/notices/65358-humanites-numeriques-etat-des-lieux-et-positionnement-de-la-recherche-francaise-dans-le-contexte-international. Dernière consultation le 12/03/2020.

Dillaerts H. et C. Boukacem-Zeghmouri (dir.), 2018. Dossier « Information scientifique et diffusion des savoirs : entre fragmentations et intermédiaires ». *Revue française des sciences de l'information et de la communication*, n° 15. Récupéré du site : https://journals.openedition.org/rfsic/4543. Dernière consultation le 12/03/2020.

Domenget J.-C., Bonaccorsi J. et V. Carayol, 2016. Introduction au dossier « Humanités Numériques et Sic ». *Revue française des sciences de*

l'information et de la communication, n° 8. Récupéré du site : http://journals.openedition.org/rfsic/1984. Dernière consultation le 12/03/2020.

Doueihi M., 2015. « Quelles Humanités Numériques ? », *Critique*, n°819-820, pp. 704-711. https://www.cairn.info/revue-critique-2015-8-page-704.htm. Dernière consultation le 12/03/2020.

Farchy J. et P. Froissart, 2010. « Le marché de l'édition scientifique, entre accès "propriétaire" et accès "libre" ». *Hermès, 57*(2), pp. 137-150.

Gitelman L. et V. Jackson, 2013. « Introduction » dans L. Gitelman (ed.), « *Raw Data* » *is an Oximoron*, Cambridge, The MIT Press, pp. 1-15.

Horn F., 2000. *L'*économie du logiciel, thèse de doctorat d'Économie industrielle, Université de Lille I. Récupérée du site : https://ori-nuxeo.univ-lille1.fr/nuxeo/site/esupversions/3916eabd-e328-4c9a-98b0-10d79d132075. Dernière consultation le 12/03/2020.

Ingold T., 2010. « L'Outil, l'esprit et la machine : Une excursion dans la philosophie de la "technologie" ». *Techniques & Culture*, n°54-55, pp. 291-311.

Jeanneret Y., 2014. *Critique de la trivialité*, Paris, Éditions Non Standard, 784 p.

Le Deuff O. et F. Clavert, 2014. « Petite histoire des humanités digitales » dans Le Deuff O. (dir.), *Le temps des humanités digitales*, Paris, Fyp éditions, pp. 15-31.

Miège B., 2004. *L'information-communication, objet de connaissance*. Bruxelles, De Boeck & INA, 248 p.

Mosco V., 2014. *To the Cloud: Big Data in a Turbulent World*. Boulder, Paradigm Publishers, 284 p.

Mounier P., 2010. « Manifeste des Digital Humanities ». *Journal des anthropologues*, n° 122-123, pp. 447-452

Mounier P., 2018. *Les Humanités Numériques*, Paris, Éditions de la Maison de l'Homme, 180 p.

Peters M., Simon, S. et P. Murphy, 2009. *Creativity and the Global Knowledge Economy*, Londres, Peter Lang, 289 p.

Rieder B., 2006. *Métatechnologies et délégation : pour un design orienté société dans l'ère du Web 2.0*. Thèse de doctorat en sciences de l'information et de la communication, Université Paris VIII Vincennes-Saint Denis.

Vinck D., 2016. *Humanités Numériques : La culture face aux nouvelles technologies*. Paris, Éditions du Cavalier Bleu, 161 p.

Waquet F., 2015. *L'ordre matériel du savoir*, Paris, CNRS Éditions, 362 p.

PARTIE III

Humanités Numériques, une question de méthodes ?

INTRODUCTION PARTIE III

Humanités Numériques, une question de méthodes ?

Dans cette troisième partie consacrée aux méthodes, la position des travaux en Sic est interrogée au sein des sciences sociales et des humanités numériques. Les travaux des chercheurs sont confrontés à des questions autant méthodologiques qu'épistémologiques. Les objets numériques sont envisagés conjointement comme des médias d'information et de communication appréhendés selon différentes dimensions, des démarches empiriques et des instrumentations de recherche, des langages formels et symboliques. La prise en compte d'enjeux socio-politiques et la mise en œuvre d'une réflexivité éthique et responsable sont soulignées par l'ensemble des auteurs.

Dans le chapitre 7, Julia Bonaccorsi et Brigitte Juanals présentent les objets et les méthodes de manière contrastée. Comme dans les autres disciplines, il n'existe pas en Sic une seule vision, une seule voix qui pourrait être consensuelle. En rendant compte de la pluralité et de l'antériorité des objets et des méthodes, les auteures mettent en relief les dialogues et les débats, dans leur dynamique et leur actualité, liées à la pluralité des approches dans la discipline. Cette pluralité prend ses racines dans l'histoire même des Sic et de sa filiation intellectuelle hybride. Elle s'appuie sur la saisie des processus d'information et de communication à partir des écritures et des dispositifs médiatiques et organisations documentaires, tout en associant les approches sociologiques de la réception des contenus médiatiques et l'analyse des usages sociaux des technologies de l'information et de la communication. L'examen des enjeux socio-politiques des mutations des formats et dispositifs de l'information et de la communication s'effectue d'une part à partir de méthodologies issues de ces ancrages

épistémologiques et, d'autre part, par des démarches empiriques menées à partir de terrains dont les périmètres sont contrastés. Les méthodologies mises en œuvre aujourd'hui tissent étroitement des démarches quantitatives et qualitatives permettant d'aborder les objets étudiés et les contextes selon plusieurs échelles. Le « numérique » constitue autant un objet d'étude qu'un outillage scientifique au service de l'analyse de la matérialité langagière et de la diffusion des savoirs.

Dans le chapitre 8, Christine Barats, Amel Fraisse, Pascal Marchand, Angeliki Monnier, Lise Verlaet tentent de caractériser la multidimensionnalité des méthodologies mobilisées en Sic. Cette multidimensionnalité est, pour ces auteurs, intrinsèquement liée au numérique lui-même, tout à la fois objet d'étude et outil de la recherche. De la collecte de l'information à la visualisation et l'interprétation des résultats, ils montrent combien la matérialité numérique a fait évoluer le rapport des chercheur·es en Sic à l'information. Ils notent au passage comment les recherches sur et avec le numérique actualisent les réflexions actuelles, tant méthodologiques qu'épistémologiques, sur les conditions de collecte des données d'une part, sur leur analyse d'autre part. Dès lors ce chapitre met l'accent sur trois questions critiques majeures relatives aux méthodologies et aux conditions de la recherche. Celles-ci sont de nature organisationnelle, économique, juridique. Et face aux problématiques posées par les usages numériques, il est souligné que les chercheurs doivent - au-delà des préoccupations épistémologiques – mobiliser trois types de compétences : technique, analytique et interprétative. Elles sont présentées à partir de cas concrets qui illustrent par ailleurs la complémentarité des méthodologies et le rôle médiateur joué par les Sic dans la construction méthodologique et l'identification des risques. De ce point de vue, les auteurs révèlent l'intérêt de la pluridisciplinarité des Sic pour saisir avec efficacité la multimodalité des données du numérique.

Dans le chapitre 9, Christine Barats, Julia Bonaccorsi, Amel Fraisse, Angeliki Monnier, Marta Severo, Lise Verlaet, abordent les Humanités numériques par les sciences humaines et sociales notamment les Sic. Les auteures investissent la question des méthodes conjointement sous l'angle des données, des outils de traitement et de visualisation ainsi que sous celui de leur interprétation. Cette approche leur permet de renouveler leurs questionnements épistémologiques, juridiques et éthiques. Les compétences techniques, analytiques et interprétatives développées en Sic nourrissent les apports critiques, tout en proposant des méthodologies et des outils éprouvés par la démarche empirique.

Les méthodologies mobilisées par les chercheurs en Sic, de la collecte de « données numériques » à la production des savoirs, questionnent les méthodes de traitement des données. Elles posent la question des enjeux et des limites des méthodologies liées au numérique, entre administration de la preuve et dimension heuristique. Elles construisent en particulier une posture réflexive afin de renforcer une démarche scientifique responsable dans leurs travaux. Dans cette perspective, les apports critiques des Sic sur les relations entre techniques, écriture, données, mémoire et savoirs sociaux, constituent des axes réflexifs puissants sur les enjeux éthiques et politiques de la recherche en SHS.

Dans le chapitre 10, Franck Cormerais aborde le développement des études digitales dans le champ des Sic autour d'une triple interrogation : celle de la manipulation des langages formels ; celle de l'apparition d'un régime de l'hyperdocument ; celle enfin d'une convergence de l'info-doc et de la communication sociale. Sa réflexion autour de l'hyperdocument l'amène à expliquer l'évolution des infrastructures réticulaires et des opérations réalisées par les technologies informationnelles associées. Il montre par ailleurs que l'invention d'un langage sert d'interface entre les langues naturelles et les langues formelles pour utiliser les possibilités computationnelles de la manipulation des symboles. Au final son analyse révèle deux éléments importants. Le premier est qu'il n'y a pas d'antinomie entre une approche statistique et une approche sémantique par adressage sémantique des données ; le second est que les Sic trouvent là une possibilité de réconcilier la division entre information et communication. Pour cet auteur les Digital Humanities valorisent une approche pluri- et transdisciplinaire dans les Sic, où l'accès à l'information scientifique et technique interroge conjointement la transformation des supports et les modes de production des connaissances. C'est donc la question de l'articulation entre l'interprétation et la computation qu'il se propose d'aborder, à partir du postulat que les études digitales constituent une opportunité pour penser une rencontre entre les humanités numériques et les Sic. Mais il milite en conclusion pour une conception élargie des études digitales, ouvrant l'hyperdocument à un processus communicationnel qui dépasse l'approche ontologique de l'informatique structurée par des métadonnées.

Brigitte Juanals
Aix-Marseille Université
Centre Norbert Elias (UMR 8562)

Éric Triquet
Avignon Université
Centre Norbert Elias (UMR 8562)

CHAPITRE VII

Pluralité et antériorité des objets et des méthodes

Julia Bonaccorsi
Université Lumière Lyon 2, UR ELICO
Brigitte Juanals
Aix-Marseille Université, Centre Norbert Elias
(UMR 8562, AMU, CNRS, EHESS, UAPV)

Résumé : Il n'existe pas en Sic une seule vision, une seule voix consensuelle. En rendant compte de la pluralité et de l'antériorité des objets et des méthodes, nous mettons en relief les dialogues et les débats, dans leur dynamique et leur actualité, liée à la pluralité des approches dans la discipline. Cette pluralité prend ses racines dans l'histoire même de la discipline et de sa filiation intellectuelle, hybride et ramifiée : depuis la saisie des processus d'information et de communication à partir des écritures et des dispositifs médiatiques et organisations documentaires, croisant celle des approches sociologiques de la réception des contenus médiatiques et l'analyse des usages sociaux des technologies de l'information et de la communication. L'examen des enjeux socio-politiques des mutations des formats et dispositifs de l'information et de la communication s'effectue d'une part à partir de méthodologies issues de ces ancrages épistémologiques, et d'autre part, par des démarches empiriques menées à partir de terrains dont les périmètres sont contrastés. Enfin, le « numérique » constitue au moins autant un objet d'étude qu'un outillage scientifique au service de l'étude de la matérialité langagière et de la diffusion des savoirs.

Mots-clés : usages sociaux, pratiques, dispositifs, technologies de la communication, culture numérique, texte, discours, litteratie, politiques, méthodes hybrides, usages des technologies de la communication, relations sociales médiatisées.

Abstract: There is no single vision, no single consensual voice. By reflecting the plurality and anteriority of objects and methods, we highlight dialogues and debates in their dynamics and their present validity, linked to the plurality of approaches in the discipline. This plurality takes its roots in the history of the discipline and in its intellectual, hybrid and branched filiation: from the capture of information and communication processes from writings and media devices and documentary organizations, crossing sociological approaches to the reception of media content and the analysis of the social uses of information and communication technologies. The discussion on socio-political challenges of changes in information and communication formats and devices is carried out, on the one hand, from methodologies derived from these epistemological anchors, and on the other hand, through empirical approaches carried out in contrasting fields. Finally, the "digital" is at least as much an object of study as a scientific tool for the study of language materiality and the dissemination of knowledge.

Keywords: social uses, practices, devices, communication technologies, digital culture, text, discourse, literacy, hybrid methods, uses of communication technologies, mediated social relations.

Ce chapitre aborde les méthodologies comme des continuums disciplinaires, en présentant certaines méthodes éprouvées en Sic à propos des objets de recherche en lien avec le numérique[1]. Sans présumer qu'il existerait des méthodologies numériques au sens propre, nous inscrivons les approches contemporaines dans la continuité de méthodologies « classiques » propres à différentes constructions des objets et des problématiques de recherche en Sic. Cette pluralité prend ses racines dans l'histoire même de la discipline et de sa filiation intellectuelle, hybride et ramifiée : depuis la saisie des processus d'information et de communication à partir des écritures et des dispositifs médiatiques et organisations documentaires, croisant celle des approches sociologiques de la réception des contenus médiatiques et l'analyse des usages sociaux des technologies de l'information et de la communication. L'examen des enjeux socio-politiques des mutations des formats et dispositifs de l'information et de la communication s'effectue d'une part à partir de méthodologies issues de ces ancrages épistémologiques et, d'autre part, par des démarches empiriques menées à partir de terrains dont les périmètres sont contrastés (musée, éducation, bibliothèque, presse, télévision, radio, Minitel, réseaux…). Enfin, le « numérique » – et ses désignations antérieures que sont par exemple les vocables successifs de « télématique », « NTIC », « TIC » – constitue au moins autant un objet d'étude qu'un outillage scientifique au service de l'étude de la matérialité langagière (traitement automatisé de la langue) et de la diffusion des savoirs (économies documentaires et sciences de l'information).

1. Voir à ce propos de manière développée : les chapitres 6 « Numérique : stratégies, dispositifs et usages » et 7 « Information, documents et écritures » de l'ouvrage collectif *Dynamiques de recherche en Sic*, publié par la Conférence permanente des directeurs.rices des unités de recherche en Sciences de l'information et de la communication, 2018, http://cpdirsic.fr/wp-content/uploads/2018/09/dynamiques-des-recherches-sic-web-180919.pdf.

Socio-sémiotique des écritures et des dispositifs médiatiques

Dans le contexte des années quatre-vingt-dix caractérisé par les discours idéologiques et promotionnels sur l'immédiateté de la transmission de l'information numérisée et l'immatérialité (« flux », « navigation », « toile ») d'une part, et d'autre part, marqué par la revendication politique de l'accès (accès aux savoirs, accès à l'information), l'analyse critique des nouvelles formes de médiatisation de l'information constitue un enjeu central pour les Sic. L'étude des usages sociaux des TIC a mis en relief cet enjeu. De manière complémentaire, la compréhension des métamorphoses médiatiques en cours et des nouveaux objets et pratiques de l'écrit, est articulée à l'explicitation des filiations avec l'histoire de la culture écrite, l'analyse des processus éditoriaux, des formats et des supports de l'écrit, l'histoire des gestes et pratiques interprétatives des textes et des savoirs. Si dans une certaine mesure les communautés en Sic qui déploient ces recherches le font prioritairement dans les secteurs de la médiation des savoirs (bibliothèque, musée) et ceux du document, le domaine des médias d'information constitue un deuxième point de vue qui va notamment participer à la compréhension des médiations (formelles, sémiotiques, matérielles, techniques) et à la redéfinition du numérique comme matérialité médiatisante, en désignant des « médias informatisés » et leurs implications socio-politiques.

Ces travaux sur les articulations entre texte, technique et usage sont particulièrement marqués par leur mise en œuvre collective (séminaires et groupes de travail locaux) puis obtention de financements nationaux comme par exemple : le Programme Télécommunications du CNRS (1998-1999) ; l'Appel d'offres BPI/DLL Écrans et réseaux : vers une transformation des rapports à l'écrit (2000-2002) ; le Programme Société de l'information du CNRS (2002-2005) ; l'Appel à Projet ANR (2006-2010). Ces recherches s'élaborent de manière interdisciplinaire croisant notamment les domaines de l'analyse de l'écriture et des médias, des sciences de la documentation, de l'ingénierie linguistique et de l'informatique. Également, ces investigations mobilisent l'enquête de terrain menée dans des espaces de pratiques pluriels et diversifiés : régies de montage radio ; formation à distance ; pratiques de présentation de consultants ; sites de diffusion littéraire ; sites de tourisme, etc.

Les méthodes développées sont nourries par la remise en question des routines méthodologiques que la discussion intra et interdisciplinaire

accompagne. Les espaces de pratiques analysés dans ces travaux constituent dans leur variété autant de moyens de formuler des questions et des hypothèses pour comprendre la complexité des objets et des pratiques concernés. « L'écriture-lecture-manipulation des textes » suppose la mise en œuvre de méthodes de description et d'explicitation des relations entre les propriétés techniques et sémiotiques des écrits numériques et les pratiques d'appropriation en tant qu'elles sont standardisées et modélisées par les médias informatisés : les chercheur.es soulignent alors l'importance d'étudier la sémiotisation des processus de communication par les médias informatisés. Rendre visibles ces processus prend appui sur différentes démarches méthodologiques ; ces démarches permettent de mettre au jour des dimensions distinctes de la matérialité des écritures numériques et de leur opérativité sociale : textualité ; interprétation ; dispositif médiatique ; strates documentaires. Ces travaux déconstruisent la notion d'interface en mettant en œuvre des approches qualitatives focalisées sur des matériaux de recherche limités et situés. Les méthodes développées visent à identifier et qualifier les signes verbaux et visuels et leur organisation signifiante dans le cadre de l'écran. Cette qualification permet ainsi de distinguer des signes ayant une fonction de manipulation et d'écriture comme le lien hypertexte (renommé textiel ou signe-passeur), d'identifier des processus de standardisation des écritures par les gabarits et formulaires (architextes) qui permettent d'alimenter les bases de données et d'éditorialiser des contenus. Le recours aux démarches d'analyse de l'énonciation échappe au pointage des fonctionnalités et de parcours de navigation, pour rendre compte des propriétés médiatisantes du numérique. Par exemple, l'étude de « sites médiateurs » de la bibliothèque numérique Gallica (200 pages de sites web renvoyant à Gallica) questionne la position métacommunicationnelle occupée par chaque site médiateur qui anticipe les usages de cette bibliothèque numérique. Cette étude caractérise l'usager-modèle de l'écriture de réseau en décrivant la nature pragmatique du texte caractéristique des sites médiateurs.

L'observation des situations interprétatives, des nouvelles gestualités et modalités de lecture, d'écriture et plus largement d'information, constitue un deuxième enjeu de connaissance et de développement de méthodologies. Irriguées par les démarches ethnographiques (comme l'observation et les entretiens d'élucidation) mais aussi par des démarches expérimentales, ces méthodes vont s'appuyer sur des protocoles explicites dans leurs limites et apports.

Par exemple, pour l'étude de sites de messagerie et de chat, vingt sites de messagerie et douze de chat constituent le « corpus » principal, et donnent lieu à un triple plan d'analyse : une description des mises en forme et des anticipations de l'usage au sein des logiciels eux-mêmes, une observation des pratiques effectives, une étude du discours des acteurs en tant que modes de qualification du jeu entre outils et pratiques.

De plus, certaines des études souhaitent saisir la lecture des médias informatisés en acte, et comprendre le jeu entre initiative et contrainte. Le dispositif d'observation défini questionne la place du signe passeur dans la lecture (ce qui est vu, reconnu) et les gestes interprétatifs mais également l'engagement corporel dans l'activité de lecture : ce dernier point a plus particulièrement donné lieu à un cadre de réflexion sur la tension entre le corps physique (postures corporelles) et le corps sémiotisé à l'écran et métaphorisé dans les architextes (pointeur de la souris, etc.). L'observation en situation (réunion de travail par exemple ou sessions de pratique organisées par les enquêteurs), s'appuie sur des consignes explicites, une analyse sémiologique des sites qui prépare les grilles d'observation, et un temps dévolu à l'élucidation avec les observé.es, ouvrant également la réflexivité sur la situation d'enquête elle-même.

Ces méthodologies ethno-sémiotiques posent les jalons d'une démarche particulièrement attentive et sensible aux effets des situations d'enquête, à la valeur de la trace d'usage et aux conditions de son interprétation sociale ; ajoutons que le « numérique » est alors caractérisé par sa continuité sémiotique avec des contextes d'espaces, d'objets, de signes divers présents là où agissent les personnes, comme c'est le cas de l'enquête sur les usages de Powerpoint par les consultants, ou encore la pratique professionnelle de montage numérique dans l'audiovisuel.

Cette double nature du média informatisé, c'est-à-dire support symbolique et artefact technique, est déconstruite de manière sémiologique, technologique (au sens de dispositif) et anthropologique. À travers ce cadre méthodologique, les questions posées par la définition de la donnée comme « trace » ou encore l'industrialisation des processus communicationnels par les médias d'écriture par excellence que sont les réseaux sociaux numériques, sont mises en exergue de manière précoce dans les recherches. De fait, les strates d'écriture (traitement informatique par le calcul de la représentation codée ; mise en forme à l'écran) et leur valeur

sociale sont ainsi étudiées depuis les années quatre-vingt-dix comme les modalités du dispositif complexe de construction, d'accès et de représentation des textes et documents que constitue le numérique.

La prise en compte de l'épaisseur documentaire en jeu dans les médias informatisés et des agencements qui configurent les conditions de la communication forme ainsi un mode d'appréhension singulier : celui-ci mobilise des méthodologies issues des champs de l'analyse des textes et du document, pour qualifier les formes inédites d'écriture, d'inscription, d'autorité et de mémoire ainsi que leur vie sociale.

De la pluralité à l'hybridation des objets et des méthodes d'analyse

Les méthodologies plurielles issues des champs de l'analyse des textes mobilisent les concepts de texte, de discours et de communication afin de décrire les objets d'étude. Ces concepts sont instables et prennent des valeurs variables en fonction des théories dans lesquelles ils sont utilisés. Dans la linguistique textuelle, le texte est considéré à la fois comme un objet abstrait, relevant de la grammaire transphrastique, et comme un objet concret, matériel, empirique, qui est le résultat toujours singulier d'un acte d'énonciation. Cette démarche permet d'appréhender conjointement les phénomènes textuels et les faits discursifs. Des phénomènes d'ordre pragmatique peuvent ainsi être analysés au niveau du texte en associant une pragmatique interne, qui reste au niveau du texte, à une pragmatique externe. Dans cette perspective, selon les analystes du discours, le discours correspond aux activités des sujets dans des contextes précis. En dépassant à la fois une analyse linguistique interne du texte et une analyse sociologique ou psychologique de son contexte, l'analyse du discours appréhende l'énonciation dans l'espace social. Le « contexte », nommé « situation de communication » par P. Charaudeau (2006), désigne l'environnement extralinguistique dans lequel se trouvent le locuteur et l'interlocuteur. On peut rapprocher le concept de « contexte » de celui de « cadre » d'E. Goffman (1991), qui correspond à l'organisation de l'expérience permettant de donner un sens aux éléments qui s'y inscrivent.

Les phénomènes de communication sont décrits par D. Bougnoux (1991) comme « la sphère des activités pragmatiques de traitement des messages entre des sujets ». La relation pragmatique entre les sujets est dite « circulaire » en référence à l'analogie entre la communication et un orchestre en train de jouer. La focalisation sur

la communication met en exergue l'importance de la relation, des médias et des médiateurs, ainsi que de la médiation. En rappelant l'opposition saussurienne entre langue et parole, Y. Winkin (1981) souligne l'apport des chercheurs de l'École de Palo Alto qui « font éclater l'opposition entre la langue essentielle et la parole accidentelle. Leur recherche ne porte ni sur la langue, ni sur la parole ; elle porte sur la communication, qui est à la fois langue et parole, compétence et exécution ». Le langage est conçu « comme une activité, et non comme le produit d'une activité ». Ils ouvrent la voie à une « linguistique de la parole » qui serait à la fois une sociolinguistique (la parole étant considérée comme une activité sociale) et une pragmatique – la parole n'étant que « l'un des multiples modes de communication mis en œuvre dans l'interaction ». La dimension pragmatique du langage est ainsi envisagée comme une pratique énonciative et comme un moyen d'action. Au sein d'une conception pragmatique du langage, elle peut s'inscrire plus particulièrement dans le cadre de la théorie des actes de langage et des énoncés dits « performatifs » (Austin, 1972), de la dimension illocutoire du langage (Searle J. R., 1972) dans ses aspects « d'acte illocutoire » et de « force illocutoire », et des travaux sur l'énonciation (Benveniste, 1966, 1974).

Par exemple, au sein d'une grande diversité des approches communicationnelles des organisations et de leurs objets d'études, les analyses portant sur l'articulation entre texte, discours et communication, se sont développées. Le langage est abordé dans ses dimensions socio-pragmatique et performative. Plusieurs travaux français ont eu recours à l'approche pragmatique du langage pour étudier le caractère illocutoire des actes de langage au sein des organisations. Dans la lignée des recherches canadiennes sur l'*organizing* comme processus de construction de sens (*sensemaking*), les travaux analysent la complexification des processus de mise en écrit des fonctionnements organisationnels. La théorie des transactions conversations-textes est associée à une conception de l'autorité considérée comme un phénomène distribué, inscrit dans des textes et actualisé dans des interactions.

De manière indissociable, il convient de prendre en considération le mode d'existence matérielle (le support matériel) du texte et son mode de diffusion – technique et humain. Cette prise en compte des aspects matériels et humains dans la caractérisation d'un texte a amené Y. Jeanneret à définir le texte comme « une configuration sémiotique empirique attestée, produite dans une pratique sociale déterminée et fixée sur un support ». Dans cette perspective, il n'existe pas d'objet

textuel observable qui soit purement linguistique. Les pratiques de lecture-écriture, toujours médiatisées, posent spécifiquement la question du rôle des techniques dans la construction de la pensée et du sens. Le texte ne peut exister sans son inscription sur un support, dont les propriétés matérielles vont influer sur la conservation, la diffusion, la médiatisation, les parcours et l'activité de lecture (J. Goody, Leroi-Gouhran). Le « cadre socio-technique » de P. Flichy (1995) s'inscrit dans une réflexion historique sur l'articulation du technique et du social.

Plus largement, les aspects sociaux des technologies de communication amènent à questionner les formes d'appropriation sociale des outils et des technologies informatiques ; ces dernières sont analysées dans les pratiques informationnelles et communicationnelles quotidiennes de la vie privée ou professionnelle. Le dépassement d'une sociologie des usages empreinte de fonctionnalisme – qui peut être présent y compris dans les problématiques centrées sur le sujet ou l'innovation sociale – invite à s'intéresser aux « usages sociaux », en référence à des pratiques liées aux caractéristiques sociales des individus, des collectifs et des institutions. L'articulation entre des cadres sociaux et culturels, des pratiques médiatiques et des dispositifs est exprimée dans le concept de « médiation sociotechnique ». Les outils et, plus largement, les dispositifs techniques portent dans leur conception l'organisation sociale et politique dans laquelle ils s'insèrent. En allant plus loin, les supports d'information sont porteurs de potentialités et de contraintes et il peut donc exister des formes de prescription par les techniques et les plateformes. Néanmoins, les recherches sur les innovations horizontales ont aussi fait apparaître les processus d'innovation par les usages dans l'univers des technologies informatiques et des télécommunications.

L'analyse des processus de redocumentarisation a fait émerger les nouvelles formes du document numérique étudiées par le collectif Pédauque (2006). Cette prescription s'opère également par les politiques et les pratiques des éditeurs d'information dans les bibliothèques ou dans l'édition scientifique.

Les innovations techniques des fabricants de matériels et de logiciels, les pratiques éditoriales des acteurs des industries culturelles et les pratiques informationnelles et communicationnelles des usagers, contribuent à construire une culture médiatique et technique et, plus largement, une culture de l'information. Cette dernière entretient des relations étroites avec la culture numérique qui recouvre, en

englobant des capacités cognitives individuelles, un ensemble de pratiques quotidiennes culturelles et sociales en interrelation avec les TIC. Les chercheurs français ont répertorié l'utilisation de nombreux syntagmes nord-américains et francophones qui font référence au développement d'une culture numérique (*digital literacy*, culture numérique, cyberculture, littératie numérique) en même temps que les cultures médiatiques et informationnelles. À cet égard, le concept de « littératie », du fait qu'il souligne un lien fort avec l'écriture, le document et le rapport entre la culture et la raison graphique, est préféré par plusieurs chercheurs à celui de « culture ». Caractérisés par leur multiplication et leur complexification, les « nouvelles littératies » liés aux environnements numériques portent des conceptions différentes. À cet égard, certains chercheurs recommandent des habiletés autant communicationnelles qu'informationnelles en vue d'une adaptation aux nouveaux environnements du web social, alors que les professionnels de l'information privilégient des compétences informationnelles. Le besoin de maîtrise d'une « culture technique » reste présent. Du fait de la convergence numérique, les diverses conceptions de la formation à l'information constituent autant de déclinaisons de l'englobante translittératie, concept qui correspond à ce constat de convergence.

La culture numérique est ainsi analysée dans le contexte des processus d'industrialisation, de numérisation et d'internationalisation de l'information. La question du rôle joué par les technologies numériques et les médias dans les modes de diffusion et les échanges est étudiée en interrelation avec les évolutions des médiations sociales, culturelles et politiques. Elle se construit à la croisée de l'offre et des stratégies des industriels, des mutations dans les pratiques éditoriales des acteurs des industries culturelles et médiatiques, ainsi que des pratiques culturelles et informationnelles des usagers des nouveaux outils et services proposés. Dans cette perspective, la culture numérique est étudiée comme un objet complexe impliquant une méthodologie pluridimensionnelle. Cette posture théorique est une manière d'affronter la plasticité conceptuelle de la médiation, dont la nature hybride – tout à la fois sociale, technique, culturelle, politique… – suscite une diversité d'approches. Les formes de la médiation sont appréhendées dans les analyses de la médiation sociotechnique et des dispositifs numériques. Elles sont constituées autant par les matériels et les services intégrant des innovations technologiques que par les modalités de leur appropriation sociale et culturelle analysées dans les usages sociaux des outils de

communication numérique. Elles sont décelables dans les politiques et les pratiques éditoriales des institutions culturelles et médiatiques, ainsi que dans les pratiques culturelles et médiatiques des publics. Elles sont en cela révélatrices des nouvelles modalités de circulation des contenus informationnels et culturels dans les sociétés contemporaines (Juanals, 2015).

Conclusion

Les méthodologies mises en œuvre aujourd'hui s'attachent à saisir de manière critique les objets complexes analysés dans ce chapitre. Pour y parvenir, elles tissent étroitement des démarches quantitatives et qualitatives permettant d'aborder les objets étudiés et les contextes selon plusieurs échelles. Le recours à des méthodologies croisées permet de dépasser une opposition entre méthodes qualitatives et quantitatives, afin de bénéficier des avantages des unes et des autres. C'est également l'une des conditions permettant non seulement de comprendre et d'interpréter, mais également d'élaborer de nouveaux instruments méthodologiques, participant de facto à la construction de nouvelles modalités de sens. À ce titre, la mobilisation croisée d'approches qualitatives (analyses de discours, lectures sémiotiques et stratégiques des dispositifs web) et quantitatives (analyses statistiques et de réseaux) s'impose pour comprendre notre objet de recherche de façon dynamique. La confrontation entre ces différentes approches ne se fait pas de façon frontale : notre souci d'étudier ensemble les contenus, les stratégies et les usages en contexte numérique nécessite un entremêlement des méthodes où les analyses statistiques nourrissent les analyses qualitatives et inversement. La compréhension fine des enjeux sociopolitiques à l'œuvre dans la médiation numérique irrigue également une réflexivité éthique et responsable à propos des méthodologies de recherche elles-mêmes.

POSITIONNEMENT

Dans l'analyse et l'interprétation des données numériques, la prudence mobilisée pour « faire parler les données » constitue une assise fondamentale pour l'enjeu scientifique qui est au cœur de la production de connaissances en sciences de l'information et de la communication. Elle suppose de situer de manière claire les objectifs : la connaissance porte non pas sur les faits sociaux et la complexité sociale auxquels donneraient accès l'analyse des objets de communication. Elle porte sur la médiation sociotechnique et les phénomènes de communication en eux-mêmes, saisis dans le temps long et par leur construction sociale. Ces constats préfigurent une position disciplinaire qui fait porter sur les méthodes de mesure et d'observation des usages du web ou des objets de communication informatisés, une distanciation nécessaire à la qualification des données en traces du social. Les méthodes développées au cours de plus de trente-cinq ans d'analyses des usages sociaux et des propriétés socio-sémiotiques des économies scripturaires du numérique questionnent directement les savoirs possibles à partir des mesures d'audience, de recueil de traces de navigation, ou encore des traces d'information, de communication ou de consommation au sein des dispositifs médiatiques.

BIBLIOGRAPHIE

Austin J., 1972. *Quand dire, c'est faire* (*How to Do Things with Words*, Cambridge, 1962), Seuil, 202 p.

Benveniste E, 1974 [1966], *Problèmes de linguistique générale, t. 1 et 2*, coll. « Bibliothèque des sciences humaines », Gallimard. 356 p . et 286 p.

Bougnoux D., 1991. *La communication par la bande*, Paris, La Découverte, 288 p.

Boure R. (dir.), 2002. *Les origines des sciences de l'information et de la communication. Regards croisés*, Villeneuve d'Ascq, Presses universitaires Septentrion, 179 p.

Charaudeau P., 2006. « La situation de communication comme lieu de conditionnement du surgissement interdiscursif » dans *Interdiscours et intertextualité et dans les médias, Tranel 44-2006*, Université de Neuchâtel, pp. 27-38.

Cotte D., 2002. « Épaisseur documentaire et numérisation : le cas des dossiers d'actualité dans la documentation de presse ». *Document numérique*, 6-1, pp. 13-28.

Escarpit R., 1976. *Théorie générale de l'information et de la communication*, Paris, Hachette Université, 220 p.

Flichy P., 1995. *L'innovation technique. Récents développements en sciences sociales. Vers une nouvelle théorie de l'innovation*. Paris, La Découverte, 255 p.

Goffman E., 1991. *Les Cadres de l'expérience* (Trad. I. Joseph *et al*.), Coll. « Le sens commun », Les Éditions de Minuit, 576 p.

Jeanneret Y., 2004. « Le procès de la numérisation de la culture », *Protée*, vol. 32, n° 2, Paris, pp. 9-18.

Jeanneret Y., 2011 [2000]. *Y a-t-il (vraiment) des technologies de l'information ?* Nouvelle édition revue et augmentée. Villeneuve d'Ascq, Presses Universitaires du Septentrion, 200 p

Jouët J., 2000. « Retour critique sur la sociologie des usages ». *Réseaux*, n°100, p. 487-521.

Juanals B., 2015. « TIC en société : pour une approche info-communicationnelle de la culture numérique » in Masselot C., Rasse P. (dir.), *Sciences, techniques et société. Recherches sur les technologies digitales*, Paris, L'Harmattan-SFSIC, pp. 25-34.

Laulan A-M., 1991. Territoires, réseaux et techniques de communication. Communication et langages, 87, p. 79-92.

Laulan A-M., RETZ, SOFEDIS, 1986. « La résistance aux systèmes d'information. » *Réseaux*, n° 19, pp. 7-24.

Le Marec J., 1989, *Dialogue ou labyrinthe, La consultation des catalogues informatisés par les usagers*, Paris, BPI/Centre Pompidou, 92 p.

Meyriat J., 1978. « De l'écrit à l'information : la notion de document et la méthodologie de l'analyse documentaire » in *Inforcom 78*, 1er Congrès de la SFSIC, Compiègne, pp. 23-32.

Moeglin P., 1994. « Votre Minitel ne demande qu'à vous rendre service » dans J.-G. Lacroix, B. Miège, G. Tremblay (dir.). *De la télématique aux autoroutes électroniques. Le grand projet reconduit*. Montréal, Québec : PUQ, Grenoble, France : PUG, pp. 163-191.

Paquienséguy F., 2006. « Entre gamme d'usages, dispositifs et personnalisation : qu'est devenu l'usage prescrit ? » Actes du Colloque international « Mutations des industries de la culture, de l'information et de la communication », [en ligne] http://icic.observatoire-omic.org/pdf/Paquienseguy3_3.pdf. Dernière consultation le 12/03/2020.

Pedauque R. T., 2006. *Le document à la lumière du numérique*. Caen, C&F éditions, 213 p.

Perriault J., 1989. *La logique de l'usage. Essai sur les machines à communiquer*. Paris, Flammarion, 253 p.

Searle J., 1972. *Les actes de langage* (*Speech Acts*, Cambridge University Press, 1969), Hermann, 264 p.

Souchier E., Jeanneret Y., Le Marec J. (dir.)., 2003. *Lire, écrire, récrire*. Paris, Bibliothèque publique d'information, 353 p. https://books.openedition.org/bibpompidou/394. Dernière consultation le 12/03/2020.

Winkin Y., 1981. *La nouvelle communication*, Paris, Le Seuil, 400 p.

CHAPITRE VIII

Regards multidimensionnels sur le numérique

Christine Barats
Université de Paris, Cerlis
Amel Fraisse
Université de Lille, GERiiCO
Pascal Marchand
Université de Toulouse, LERASS
Angeliki Monnier
Université de Lorraine, Crem
Lise Verlaet
Université Paul-Valéry Montpellier 3, LERASS

Résumé : Ce chapitre vise à caractériser la multidimensionnalité des méthodologies mobilisées en Sic à partir d'enjeux de recherche qui apparaissent spécifiques à la discipline, ou du moins qui font l'objet d'une réflexivité remarquable dans différentes productions et projets collectifs. La pluridisciplinarité des Sic constitue également un véritable atout pour se saisir du numérique. En prenant appui sur des exemples de collaborations pluridisciplinaires, nous rendons compte de certains phénomènes émergeant des confrontations et articulations avec d'autres disciplines sur les méthodologies. Ce chapitre met en particulier l'accent sur trois questions critiques majeures liées aux méthodologies et aux conditions de la recherche (organisationnelle, économique, juridique).

Mots-clés : méthodologies, corpus, compétences numériques, projets de recherche Sic.

Abstract : This chapter aims at characterizing the multidimensionality of the methodologies mobilized in information and communication sciences from research issues that appear specific to the discipline, or at least that are the object of a remarkable reflexivity in different productions and collective projects. The multi-disciplinary nature of information and communication sciences is also a real asset when it comes to dealing with digital technology. Using examples of multidisciplinary collaborations, we report on some of the phenomena that emerge

from confrontations and articulations with other disciplines on methodologies. In particular, this chapter focuses on three major critical issues related to methodologies and research conditions (organizational, economic, legal).

Keywords: methodologies, corpora, digital skills, CIS research projects

La multidimentionnalité des méthodologies est intrinsèquement liée au numérique lui-même, tout à la fois un objet d'étude en tant que technologie de l'information et de la communication, un domaine de recherche du fait de l'omniprésence des dispositifs socio-numériques et un outil désormais incontournable dans les projets de recherche. Néanmoins, sous certains aspects, le numérique n'est qu'un nouvel objet-domaine-outil parmi d'autres étudiés en Sic[1] et que notre discipline sait circonvenir par les méthodes qu'elle emploie ou développe. Car les phénomènes ne sont pas forcément nouveaux, mais changent d'échelle. Cela étant, le numérique a métamorphosé bon nombre d'objets d'étude usuels en Sic, constituant par là même un terrain fertile de réflexion épistémologique, scientifique et professionnelle.

En effet, alors que les objets de recherches (les pratiques, les contenus, les dispositifs, les formes de médiation ou des enjeux communicationnels) sont de plus en plus souvent numériques, les conditions de collecte, d'analyse, d'interprétation, de visualisation et d'archivage de sources numériques de différentes natures constituent un nouvel enjeu pour les recherches en SHS.

L'une des idées reçues quant aux recherches sur et avec le numérique réside dans l'apparente accessibilité des sources et invite ce faisant à être questionnée, car les sources et données numériques ne sont pas stables dans le temps, ce que l'on nomme la *vélocité* des données numériques en référence à la vitesse à laquelle elles sont générées, déplacées, voire supprimées. De plus, force est de constater que le temps de la recherche coïncide rarement avec celui des dispositifs et des pratiques inhérentes, même si des approches ayant recours à l'automatisation de certaines procédures permettent - dans des délais qui se raccourcissent- de livrer des premiers résultats sur des tendances ou orientent vers des pistes de recherche. Les premiers résultats liés à l'exploitation des données sont autant de valeurs sur lesquelles le chercheur peut s'appuyer pour poursuivre les analyses

1. Voir *Dynamiques des recherches en sciences de l'information et de la communication*, ouvrage édité par la Conférence permanente des directeurs·trices des unités de recherche en sciences de l'information et de la communication, juillet 2018.

complémentaires de son protocole de recherche à des fins de validation, et/ou éventuellement de le faire évoluer.

De plus, ces sources et données numériques se caractérisent par une hétérogénéité et une multimodalité (traces, signes, données textuelles, images, sons, architextes, codes, mesures…) qui affectent les notions de données, de corpus, de pratiques ou d'observables et d'archives. Que ce soit au regard des pratiques, des contenus, des dispositifs, des formes de médiation ou des enjeux communicationnels, les recherches sur et avec le numérique impliquent ainsi des regards multidimensionnels afin de saisir la complexité des données collectées et de définir des niveaux d'observation et de collecte qui circonscrivent le périmètre d'observation en fonction des besoins de la recherche. Ces regards multidimensionnels facilitent la sélection des données, ils étayent en amont l'élaboration du protocole de recherche, en éclairant les critères de sélection et de collecte, sans mésestimer les contraintes techniques de collecte et de stockage. Il s'agit, de plus, de distinguer, et parfois d'associer, des données nativement numériques (issues par exemple des réseaux socionumériques, contenus web), et des données numérisées (archives antérieures à l'ère numérique ou supports produits sans intention de diffusion numérique), tout autant que de tenir compte de l'instabilité des dispositifs et des contenus dans la mise en œuvre du protocole. L'analyse des données requiert également d'articuler les approches quantitative et qualitative afin de mettre au jour et examiner leurs conditions de production.

Les recherches sur et avec le numérique actualisent les réflexions tant méthodologiques qu'épistémologiques, que ce soit au regard des conditions de collecte des données ou de leur analyse. Les Sic contribuent ainsi à rendre intelligibles les phénomènes information-communicationnels complexes (Verlaet, 2015). En effet, les sources et données numériques nécessitent de mobiliser des approches multidimensionnelles afin de saisir leurs spécificités et mettre en œuvre un protocole pertinent au regard de la problématique. Il s'agit de surcroît de veiller à la lisibilité du protocole dans la durée, ce qui suppose une documentation des conditions de collecte, de stockage et d'analyse.

Finalement, l'analyse des corpus issus du web (archives en ligne, réseaux socionumériques, sites, forums et débats institutionnels ou non) doit mobiliser au moins trois types de compétences: technique, analytique et interprétative.

Des compétences techniques

Que la recherche ait recours ou non à des bases de données, la connaissance des conditions de production des données ainsi que celle de leur écosystème est une condition *sine qua non* pour la mise en œuvre du protocole de recherche tout autant que pour l'interprétation des résultats. Dans le cas de collecte de données structurées et ouvertes à partir de bases de données nationales ou internationales (INA, BNF, Sudoc, WorldCat, ...), ou bien d'utilisation de données numérisées du type Google books ou collecte de données non structurées issues du Web et des réseaux sociaux, il s'agit également de connaître les critères d'archivage (stabilité ou non des collections dans le temps, fréquence de collecte, périmètre de collecte...) et de numérisation (perte ou non de données...) qui conditionnent le recueil des données et peuvent restreindre ou orienter la problématique de la recherche. Ces éléments sont primordiaux pour la constitution de corpus, qui rappelons-le, sont des construits scientifiques élaborés en fonction des perspectives de recherche.

De récents travaux de recherche en Sic mettent en place aujourd'hui des nouveaux paradigmes de collecte et de structuration de données sur une échelle large, transnationale et multilingue, afin d'aboutir à un partage d'archives, de ressources et de connaissances dans un contexte mondialisé. Ces travaux proposent d'allier, de manière symbiotique, intelligence collective mobilisant des ressources humaines pour la collecte et l'annotation des données (crowdsourcing) et intelligence artificielle mobilisant des ressources automatiques pour l'analyse, l'extraction et la visualisation de l'information.

Ainsi, les projets DENIM (financé par le ministère de la Culture) et Mémo-Mines (projet ANR-16-CE38-0001-02) partagent l'objectif d'améliorer la visibilité des patrimoines textile et minier des Hauts-de-France en se basant sur des technologies du web sémantique. Il en va de même pour le projet interdisciplinaire et international ROSETTA : tout comme la pierre de Rosette a permis le déchiffrement des hiéroglyphes grâce à la présence de la traduction grecque, le projet ROSETTA vise à préserver des patrimoines culturels et linguistiques menacés de disparition et à les développer grâce à l'existence de supports langagiers numérisés dans les collections des bibliothèques et archives nationales. Une collaboration entre l'université de Lille et l'université de Stanford vise ainsi à soutenir la représentation numérique de la diversité linguistique et culturelle des savoirs en mettant en place des approches numériques permettant une collecte

exhaustive, un traitement automatique et une visualisation interactive des informations bibliographiques multilingues relatives à une même œuvre transnationale afin de mieux analyser et comprendre sa réception locale à travers les différents contextes culturels.

Et, ici comme ailleurs, la notion de corpus est primordiale: « Il n'y a […] pas de bon corpus *a priori*, mais des corpus cohérents ou non avec des hypothèses et des problématiques » (Bonnafous, Krieg-Planque, 2013 : 231). Le corpus s'ancre dans un terrain, c'est-à-dire « les situations ou les lieux empiriques où le chercheur va collecter des données » (Boutet, 2002 : 568). La multidimensionnalité inhérente au numérique rejaillit nécessairement sur la constitution du corpus tant les terrains offerts par le numérique sont pluriels. Le numérique peut être considéré comme une « œuvre ouverte » au sein de laquelle les « 5 V » des données massives (*Volume, Vélocité, Variété, Véracité* et *Valeur*) réinterrogent les notions d'exhaustivité, de représentativité, d'homogénéité des corpus, tout comme leur traitement et leur interprétation (Anquetil *et al.*, 2019).

Que la collecte soit faite en ligne ou qu'elle s'adosse à des bases de données (*Factiva, Europress*, INA, dépôt légal…), le chercheur ne délègue pas à la machine la constitution de son corpus: il le construit, le contrôle, le justifie, prend en compte ses caractéristiques et ses évolutions.

Pour autant, la compétence technique peut nécessiter un rapport au langage informatique d'encodage des données numériques, que ce soit dans les bases de données ou dans les hypertextes, qui permet l'extraction des données utiles à la recherche, leur formatage en fonction du type de traitement envisagé (distinguant par exemple les données et les métadonnées) et leur « nettoyage » pour neutraliser les aléas de saisie, de numérisation ou d'encodage qui peuvent être à l'origine de biais.

De nombreuses recherches en Sic ont ainsi mis l'accent sur le temps long nécessaire, que la collecte soit supervisée ou non par les chercheur·es.

L'évolution des outils d'aide à la saisie numérique (scanners à reconnaissance optique des caractères, reconnaissance vocale…), mais également l'accès à des bases de données toujours plus riches, amènent à constituer des corpus de plus en plus grands. En quarante ans, de 60000 occurrences des « tracts de mai 1968 » on est passé, par

exemple, à 238000000 occurrences du « Cablegate »[2]. Les Sic doivent pouvoir s'en emparer et produire, dans des temps raisonnables, des analyses qui prennent en compte les intérêts et limites de tels volumes de données.

Citons, pour illustrer, la recherche menée sur la circulation de l'information sur *Twitter* qui a fait appel à une entreprise pour collecter les *tweets* qui pointaient vers les 31 médias examinés et a exigé un temps long, presque un an, pour éliminer les doublons et les *tweets* hors problématique et ne retenir *in fine* que 17 millions de *tweets* par rapport aux 38 millions collectés (Mercier, Pignard-Cheynel, 2018 : 15).

Ces questions, qui mobilisent des compétences en sciences des données, peuvent nécessiter la mise en place de collaborations interdisciplinaires. Ainsi, le projet LisTIC[3] se donne à comprendre comment les usages des réseaux socionumériques et des applications de *micro-blogging* développés à l'aide des smartphones s'inscrivent dans l'écosystème médiatique des e-participations. Il s'agit d'explorer les usages contemporains des téléphones mobiles et des applications développées par les principaux médias sociaux dans le but de mieux cerner leurs effets sur la composition et la morphologie des réseaux relationnels, sur les sociabilités tissées au travail et sur les formes médiatisées de la participation politique, que ce soit en France durant les élections présidentielles ou dans le cadre des mouvements sociaux actuels au Brésil. Pour cartographier ces pratiques, le projet propose de développer les méthodologies des humanités numériques en exploitant les ressources offertes par les smartphones afin d'en documenter les usages, notamment les formes nomades de la participation aux applications mobiles des réseaux socionumériques. C'est un projet interdisciplinaire impliquant des chercheurs et unités de recherche en sociologie, en sciences de l'information et de la communication et en informatique[4].

Enfin, il est de moins en moins rare que des chercheurs en Sic développent les compétences en sciences de données nécessaires à ce type d'opérations. La question peut se poser de la formation des

2. Il s'agit des 251 287 câbles de la diplomatie américaine mis en ligne par le site WikiLeaks, et qui constituent les communications entre les ambassades américaines du monde entier et l'administration de Washington.

3. Projet financé par l'Agence Nationale de la Recherche (ANR-16-CE26-0014-01) avec le soutien du GDRI Web Science.

4. Les différentes enquêtes réalisées dans le cadre de ce projet, à l'aide des méthodes des humanités numériques, ont été mises en place à l'aide de la plateforme OSIRIM.

chercheurs en Humanités numériques aux bases et rudiments des sciences de données leur permettant *a minima* de dialoguer avec les sciences et techniques numériques et, pour certains, de prendre en charge l'extraction et le formatage des données. Se pose également la question de leur stockage et archivage ainsi que leur pérennité, ce qui suppose des moyens techniques et humains qu'il s'agit d'anticiper, dans un contexte juridique de plus en plus contraignant et variant selon les pays. Comparer l'expression des discours de haine en ligne dans divers pays européens se heurte, par exemple, à des législations nationales qui vont autoriser, ou non, ou différemment, l'extraction, le stockage, le codage des métadonnées et la publication des résultats… Et on ne gagnera pas en simplicité si l'origine des discours provient de dispositifs développés et hébergés dans des espaces et des environnements juridiques extra-européens. Si la TIGR Huma-Num entend dans une certaine mesure répondre aux problématiques de stockage et d'archivage, les questions juridiques restent à démêler lorsque les recherches s'appliquent à des contextes internationaux.

Des compétences analytiques

Les dispositifs numériques se caractérisent par une extrême variété et hétérogénéité des données (contenus, pratiques, flux) ainsi que, comme nous venons de le voir, par leur volume. Le statut même de la donnée en Sic est posé, et notamment des données massives. Il peut s'agir de travailler les données elles-mêmes (compétences en sciences des données), de travailler sur les données (compétences en statistiques textuelles, détection de communautés, web sémantique), de travailler avec les données (compétences en analyse sémio-pragmatique, ethnographie en ligne) ou de travailler à propos des données (compétences en épistémologie, analyse réflexive et critique).

Les méthodologies mises en œuvre en Sic mettent également l'accent sur ce que change le numérique par rapport à d'autres dispositifs de communication. En l'occurrence, l'instabilité des dispositifs – objet de modifications régulières – implique de documenter la collecte des données afin de garantir leur lisibilité ainsi que leur interprétation dans la durée, en l'occurrence à court et à moyen terme. Citons, dans une perspective diachronique, le cas de dispositifs qui ont marqué les débuts du développement du web comme *Myspace* ou plus généralement l'importance des blogs qui exigent aujourd'hui pour comprendre leur écosystème d'avoir sauvegardé des traces de leur environnement pour situer les éléments et données collectés (captures d'écran et documentation des données sauvegardées). Citons

également, dans une perspective synchronique, les modifications régulières de l'interface de *Facebook* ou de *Twitter* qui nécessitent de documenter la collecte, à la fois au niveau du dispositif examiné et également du type de données collectées et des niveaux de collecte : à quelle période s'effectue la collecte ? À partir de quel terminal de consultation (ordinateur, tablette, smartphone…)? Quelles sont les données collectées (textuelles, sonores, images, vidéos) et quelles sont celles qui ne sont pas collectées (publicités, images animées, gif…) ? Quelle est la place accordée aux architextes et aux codes sources dans la collecte? Comment garder des traces de l'écosystème observé, des fonctionnalités et mesures qui y sont associées? Rappelons néanmoins que la BnF assure l'archivage du web depuis 2006 dans le cadre de la mission de dépôt légal à partir d'une sélection de sites, mais ne couvre pas l'ensemble du web.

Dans le cadre du projet APPEL (*Analyse Pluridisciplinaire du Pétitionnement En Ligne - Pluridisciplinary Analysis of On Line Petitioning*)[5], un groupe de chercheurs a eu accès aux données de l'un des principaux sites francophones de pétitions en ligne (lapetition.be), soit plus de 12000 pétitions ayant récolté au total 3,25 millions de signatures sur la période comprise entre le 31 octobre 2006 et le 12 février 2015. La richesse des données (inscrites par les internautes et également générées par le dispositif) de même que leur hétérogénéité (données alphanumériques, nombre de signataires, rubriques et formulaires du site…) ont conduit à articuler différentes approches: qualitatives (analyses d'entretiens avec le concepteur) et quantitatives (outils de statistique textuelle) afin de prêter attention aux caractéristiques du pétitionnement en ligne, ainsi qu'aux contraintes et incitations induites par le dispositif qui affectent la dynamique de pétitionnement. L'entretien s'est avéré essentiel pour apprécier la stabilité du dispositif dans le temps et ne pas biaiser les interprétations des données collectées. Il a ainsi permis de mesurer la stabilité du dispositif par rapport à d'autres plateformes de pétitionnement, comme par exemple change.org dont les rubriques de classement des pétitions ne sont pas stables dans le temps. L'entretien a également permis de prendre connaissance des choix du concepteur: pas de conseils en ligne pour les porteurs de pétition, contrairement à d'autres plateformes, nombre stable de rubriques visant à classer les pétitions…

5. Projet financé par l'Agence Nationale de la Recherche.

Le croisement des données (textes d'appel, nombres de signataires, rubriques de classement des pétitions…) a permis de mettre au jour le poids du dispositif dans l'acte de pétitionnement. Par exemple, l'affichage en page d'accueil des dernières pétitions mises en ligne ou les plus signées augmente le nombre de signatures, confirmant de nombreux travaux sur l'importance de la mise en visibilité des mesures et leur dimension amplificatrice. Le parti pris a été de tenir compte à la fois du dispositif et de son design, ainsi que des différents types de données, en l'occurrence les architextes du site (rubriques, formulaires de dépôt ou de signature d'une pétition…), les données inscrites par les internautes et celles générées par le site (nombre de signataires, les plus signées…). Les données collectées ont éclairé les incitations et contraintes induites par le dispositif. Par exemple, l'analyse de l'incidence des rubriques sur l'acte de pétitionnement a montré la diversité des modalités de leur appropriation, à savoir leur appropriation distancée ou non, et parfois leur détournement par des réappropriations ludiques… Si le concepteur a anticipé les pratiques de détournement en proposant une rubrique « Humour/insolite », toutes les rubriques proposées ont fait l'objet de formes de détournement. L'analyse de la rubrique « Humour/insolite » a montré la présence de pratiques de sociabilités familiales et professionnelles (pétitions humoristiques dans le cadre familial ou professionnel), indice d'une réappropriation de l'acte de pétitionnement et de son articulation à d'autres formes de sociabilités. Les textes d'appel ainsi que leurs commentaires ont permis de constituer différents corpus afin de prêter attention à leurs caractéristiques et traits saillants. L'analyse des textes et des commentaires a ainsi montré l'importance de relais externes (médias, associations, pages Facebook…) dans la dynamique de pétitionnement, soulignant la dimension transmédiatique de ces dispositifs ainsi que l'articulation de pratiques en ligne et hors ligne. L'examen des particularités des écrits du web a par ailleurs souligné leur circulation (émoticônes) ainsi que leur dimension techno-langagière. Si le volume des données conduit à repenser la notion de corpus, il invite également à croiser les données afin de constituer des corpus *ad hoc*, en lien avec les hypothèses de recherche et à mobiliser des outils d'analyse statistique de données textuelles, compte tenu du volume des données. Les outils ont facilité la mise en évidence des caractéristiques discursives des textes de pétitionnement (formes d'adresse, autres relais web) et de leurs commentaires qui éclairent les raisons pour lesquelles une pétition est signée. Cependant, il s'agit d'effectuer systématiquement des retours au texte, afin de vérifier

l'interprétation des données et tenir compte de l'hétérogénéité des graphies et de leur contexte. Si les outils sont automatiques, l'interprétation ne l'est pas et nécessite un retour aux textes et contextes pour garantir les résultats.

Ces orientations méthodologiques multidimensionnelles mobilisées en Sic permettent de situer et de contextualiser les données et s'avèrent déterminantes pour le choix des outils d'analyse ainsi que pour l'interprétation des résultats. Il s'agit de tenir compte des niveaux de collecte (en surface vs en profondeur avec les codes…) ainsi que des angles morts de la recherche, c'est-à-dire ce que le chercheur ne prend pas en compte ou ne peut pas prendre en compte (exemple : Facebook, collecte à un temps "t", *quid* du fil d'actualité, des publicités, des gifs… etc.).

Le volume des données numériques et leur hétérogénéité (textes alphanumériques, ponctuation, smileys, gifs, mesures…) impliquent de faire des choix et de ne pas mésestimer le temps de collecte et de préparation des données.

La compétence analytique requiert la mise en place de méthodes diverses selon des approches et des sensibilités qui s'inscrivent dans des traditions de recherche plurielles. L'objectif est alors de faire converger des protocoles susceptibles de livrer un aperçu exhaustif de l'appropriation contemporaine de technologies émergentes et de la manière dont elles infléchissent, voire renouvellent, nos structures sociales et relationnelles. Pour relever cet ambitieux défi scientifique, des méthodologies innovantes sont développées en lien étroit avec les méthodes classiques des SHS et leurs acquis théoriques pour favoriser au maximum la cumulativité des connaissances. Une ethnographie en ligne peut être associée à des outils de collecte de traces d'utilisation et de cartographie des communautés en ligne (par exemple au sein de *Twitter* et *Facebook*), de collecte des messages et des commentaires des internautes afin de produire une analyse graphique des débats en ligne. Cette étape pourra, éventuellement mais pas obligatoirement, recourir à des instrumentations plus ou moins complexes. Les méthodes actuelles permettent de décrire la structure de données en se reposant sur leurs contextes (cartographies, classifications automatiques, profilages, extractions…).

La visualisation des données après traitement automatique peut revêtir un aspect de plus en plus attractif et diffusable (graphes de relations, nuages de mots, graphes dynamiques…), mais toute représentation est toujours une compression avec perte d'information et l'analyste doit en maîtriser les procédures pour en garantir l'interprétation et éviter des

interprétations intuitives et orientées. Le risque du simple comptage de mots conduisant à une interprétation décontextualisée doit être considéré. Les grandes fréquences ont leur importance, car la redondance génère du sens, mais elles ne suffisent pas à décrire un corpus.

Ce sont alors des regards différents qui sont mobilisés sur le corpus, de la matérialité lexicale aux effets sémantiques, en passant par la catégorisation syntaxique, les marqueurs stylistiques, les catégorisations du contenu… La passation d'entretiens auprès de militants ou d'internautes particulièrement actifs pourra rendre compte de la manière dont ils s'approprient ces outils afin de partager des informations, débattre et tisser des sociabilités numériques.

Des compétences interprétatives et critiques

Les objets traditionnellement étudiés en Sic mobilisent des approches multidimensionnelles. Par exemple, les travaux qui examinent des textes et contenus, prêtent attention à leurs contextes et à leurs conditions de production afin de les appréhender et de les analyser dans leur écosystème. Ce parti pris s'avère essentiel pour examiner les productions numériques, qu'elles soient numérisées ou nativement numériques, car il s'agit de tenir compte de leurs modalités de production et de leur environnement afin de les contextualiser, que ce soit en amont de la recherche pour définir un protocole ad hoc en lien avec la problématique de la recherche, ou que ce soit au moment de la collecte des données pour circonscrire le périmètre et les modalités de collecte ou bien en aval, lors de leur analyse, afin de garantir l'interprétation des résultats.

De la collecte de l'information à la visualisation et l'interprétation des résultats, la matérialité numérique a fait évoluer le rapport des chercheur·es en Sic à l'information, en décloisonnant notamment les champs des sciences de l'information et des sciences de la communication, en les rassemblant à partir de la notion de « donnée ». Ainsi, l'usage de nouveaux vocables comme "data" ou le "Big Data" est entré dans les pratiques, aussi bien comme objet de recherche (en témoignent les travaux portant sur l'Open Data), que comme matériau de recherche (collecte, analyse et traitement de données nativement numériques ou numérisées).

Aborder des corpus numériques sous l'angle de leur contenu amène à analyser les traces d'opérations matérielles que l'on peut appréhender à des niveaux divers. Il s'agit de « pister la *traçabilité* des dires rencontrés à propos de certains événements au fil de leur

voyage » (Moirand, 2007 : 1). Le texte est alors la trace mémorielle d'un événement discursif intelligible dont il faut établir la définition pour décrire le corpus. Cette analyse de la trace peut être assimilée à une démarche darwinienne qui conduit à constituer patiemment des corpus de tailles variables mais de plus en plus importantes pour étudier les traces qu'il contiennent et en inférer les conditions de production. On se situe alors dans une démarche plus empirique qu'hypothético-déductive, mais plus heuristique qu'herméneutique.

La numérisation des *Archives parlementaires de la Révolution française*, par exemple, est un travail mené en collaboration entre les bibliothèques de l'Université de Stanford et la Bibliothèque nationale de France (BnF). Les analyses textométriques en cours de ce corpus couvrant les années 1787 à 1794 (plus de 43,4 millions de mots) montrent que la numérisation est imparfaite et incitent à une amélioration des procédures de constitution de l'archive elle-même. Néanmoins, elles permettent de retrouver des spécificités et des ruptures bien connues des spécialistes de la Révolution française: on retrouve aisément les changements d'assemblées, de la *Constituante* à la *Législative* puis à la *Conventionnelle* qui valident la pertinence de l'approche. On observe également des faits nouveaux et on peut apporter des éclairages sur certaines controverses fameuses.

Les travaux qui prêtent attention aux usages examinent les changements d'échelle, l'articulation entre « en ligne » et « hors ligne », le poids des mesures et des algorithmes, etc. La nature même des écrits du web amène à poser des problématiques spécifiques vis-à-vis d'autres SHS. Les notions d'anonymat et de pseudonymat, ou plus largement d'identités numériques (Merzeau, 2009), de même que l'inégalité d'accès aux dispositifs, posent la question de la représentativité des observations et peuvent empêcher l'interprétation et la généralisation des résultats en termes sociocatégoriels ou démographiques. L'évolution de l'expression en ligne (oral-écrit, style SMS, techno-langage, émoticônes, marqueurs évaluatifs…) amène à analyser des données hétérogènes en grand nombre en se gardant de réduire les résultats et de produire des artefacts.

Le projet M-PHASIS (*Migration and Patterns of Hate Speech in Social Media. A cross-cultural perspective*)[6], cherche à étudier les discours haineux contre les migrants dans les médias socio-numériques. Il vise à mieux appréhender la prévalence et l'émergence de ceux-ci dans les

6. Financé par l'Agence nationale de la recherche et la Deutsche Forschungsgemeinschaft (ANR-DFG, portant la référence ANR-18-FRAL-0005).

contenus générés par les utilisateurs en ligne, en France et en Allemagne. Par ailleurs, le projet aspire à bénéficier des apports de l'Informatique en matière de traitement automatique des langues, avec l'objectif – entres autres éléments – de produire une application logicielle détectant et bloquant automatiquement les commentaires haineux en ligne. Pour cela, il s'appuie sur la technologie du Deep Learning (« apprentissage profond »), selon laquelle des données étiquetées manuellement comme exemples « haineux » permettent à un ordinateur d'apprendre à étiqueter d'autres données similaires de manière autonome.

Le projet réunit une équipe de chercheurs en sciences de l'information et de la communication, en sciences du langage et en sciences de l'informatique. La collaboration s'appuie sur la complémentarité des compétences et profite d'un vrai dialogue et d'une réelle volonté d'échanges entre les membres ; mais elle se confronte également aux pratiques et préoccupations disciplinaires qui façonnent des regards et des attentes parfois diversifiées. En sciences de l'informatique, la demande consiste surtout à la constitution d'un « lexique de haine » afin d'améliorer, en termes de validité et de fiabilité, l'apprentissage de la détection automatique du discours haineux par la machine, dont les zones d'ombre sont explicitement admises par les ingénieurs. La question de la circulation des messages intéresse davantage les sciences de l'information et de la communication, tandis que, de manière plus générale, Sic et sciences du langage portent leur attention au-delà du paradigme proto-lexical, en interrogeant les paramètres pré-discursifs (implicites, représentations partagées) et post-discursifs (orientation actionnelle) des propos observés. Le traitement des données langagières reflète aussi les divergences disciplinaires. En sciences de l'informatique (comme en linguistique informatique, par ailleurs), les corpus, une fois collectés, sont nettoyés d'éléments formels, e.g. ceux de ponctuation, et lemmatisés. En sciences du langage, en revanche, et notamment pour les chercheurs qui adoptent des approches qualitatives, les ponctuants (points de suspension, points d'exclamation, etc.) et autres aspects formels (e.g. usage des majuscules) sont signifiants.

L'absence d'une « norme de référence » concernant la définition du discours de haine constitue un obstacle de taille dans la spécification du protocole de la recherche. En linguistique informatique, le recours à des listes préétablies de mots « haineux » reste la pratique la plus courante. Mais ces listes arrivent difficilement à capter les formes les plus subtiles – et donc implicites – de haine (sarcasmes, euphémismes, stéréotypes, références contextuelles, par exemple historiques ou liées à l'actualité), ainsi que les « stratégies de masquage » mises en place

par les internautes (jeux avec l'orthographe, crypto-langages réservés aux initiés, etc.). Dire ce qui relève ou pas de la « haine » s'avère alors difficile, aussi bien pour les annotateurs humains que pour les machines (algorithmes). D'autant plus que le contenu en ligne à analyser est vaste.

Dans ce cadre, une posture « holistique » semble la plus apte à mettre en évidence les processus multidimensionnels, interactionnels et connectifs, au sein desquels émergent les contenus haineux en ligne. Il s'agit à la fois d'analyser les messages (traces langagières), les dispositifs (affordances des plateformes de réseaux socio-numériques, rôle des algorithmes, politiques de modération), les usages et les profils des internautes (genre, localisation géographique, recours à l'anonymat, etc.), les logiques de mise en scène et de visibilité qui sous-tendent les pratiques (*playfulness*, authenticité de soi, appartenance, *empowerment*, etc.), ainsi enfin que les contextes juridiques, historiques, socio-culturels, etc. qui définissent le sens du discours de haine et les modalités de sa sanction. Au-delà des intentionnalités individuelles, émergent ainsi un macro-discours thématisé ainsi qu'un phénomène social d'une « haine » au sens large du terme, à savoir hostilité ambiante et « motile », avec des manifestations verbales et non verbales extrêmement diverses, et une violence scalaire allant de propos méprisants à des incitations à des actes criminels.

Par leur multidimensionnalité, les Sic sont très bien placées pour répondre aux besoins d'une telle recherche. Cela implique des choix qui ne sont possibles qu'à condition de connaître leur environnement et écosystème. Il faut donc documenter les conditions de collecte des données ainsi que leur périmètre afin de tenir compte du design du dispositif, de ses fonctionnalités et des spécificités des interfaces à un temps "t", celui de la collecte.

Dans cette perspective, les travaux en Sic articulent différentes approches et ont recours à des entretiens avec les concepteurs des dispositifs et/ou font appel à différents types de documents qui permettent de contextualiser et de situer dans leur écosystème les données collectées. La dimension éphémère des données et leur vélocité exigent également un protocole de collecte et de stockage afin de garantir leur accès et leur lisibilité sur la moyenne durée.

Cette compétence interprétative reste la plus attendue des sciences humaines et sociales, car elle fonde leur utilité, du moins aux yeux de disciplines relevant des sciences-technologie-médecine (STM). La pluridisciplinarité des Sic permet de saisir la multimodalité des données du numérique: pratiques d'utilisateurs, textes et signes,

images, sons, flux, mesures, etc. Il s'agit avant tout, en puisant dans les théories et modèles des sciences humaines et sociales et en mobilisant les connaissances disponibles sur l'environnement social de la recherche, d'articuler des textes, des cotextes et des contextes afin de situer les pratiques et les dispositifs, sans les naturaliser, ni les réifier.

Les cas concrets ici présentés permettent d'illustrer les logiques de complémentarité des méthodologies, le rôle médiateur joué par les Sic dans la construction méthodologique et l'identification des risques.

La collecte de données et la construction automatique ou semi-automatique des connaissances et des liens sémantiques qui peuvent les relier est un défi majeur des systèmes d'organisation de connaissances aujourd'hui. Ces derniers doivent mettre en place des modèles d'encodage et de balisage facilitant l'interopérabilité multilingues et multiculturelles. L'interopérabilité des systèmes d'organisation des connaissances, et de corpus de données d'une manière générale, facilitera l'émergence de nouveaux questionnements et de nouvelles problématiques.

POSITIONNEMENT

Face aux problématiques posées par les usages numériques, les chercheurs doivent associer aux préoccupations épistémologiques usuelles trois compétences: technique, analytique et interprétative.

Inhérente au numérique est la question des données, dont les Sic doivent éminemment se saisir pour faire valoir leur vision multidimensionnelle du numérique, lequel soulève bien des écueils que nous savons identifier et circonscrire. À ce titre et du fait d'être Sic, les chercheur·es de notre discipline peuvent très largement concourir aux dialogues entre les sciences humaines et sociales et les sciences du numérique. Mais pour que ceux-ci restent féconds, il est nécessaire que les chercheur·es en Sic puissent prendre connaissance des transformations des outils de recueil, des approches et modèles de traitement de données, des problématiques qui les sous-tendent, des orientations épistémologiques qu'ils induisent; et cela ne peut se faire qu'à partir d'une démarche de formation continue, leur donnant la possibilité de suivre et de comprendre les évolutions rapides relatives aux sciences des données.

Plusieurs projets de logiciels *open source* portés par des laboratoires et chercheurs en Sic sont remarquables et auront à jouer un rôle croissant dans les années à venir. Ces derniers dénotent de l'étonnante polyvalence de notre discipline pour produire des outils collaboratifs performants et librement accessibles servant tout à la fois les entreprises scientifiques et citoyennes. Inutile de souligner l'importance des ressources humaines et financières qui doivent être mobilisées pour répondre à ces nouveaux enjeux.

BIBLIOGRAPHIE

Anquetil S., Duteil-Mougel C., Lloveria F. (dir), *Le sens des données. Statut du corpus et herméneutique à l'aune des humanités numériques*, L'Harmattan, 2019.

Barats C., Fiala P., Leblanc JM., « Approches textométriques du web: corpus et outils », In Barats C. (dir), *Manuel d'analyse du web*, A. Colin, 2016, 107-132.

Bolka L., M., Thiault F., Després-Lonnet, « Déplacer les frontières socioculturelles pour faciliter l'appropriation d'un dispositif numérique. Expérimentation d'un processus de co-conception », In Hachoui H., Bouhai N. et Saleh I., (dir), *Frontières numériques et savoir*), Paris, L'Harmattan, 2016, 95-118.

Bonaccorsi J., « Approches sémiologiques du web », In Barats C. (dir), *Manuel d'analyse du web*, A. Colin, 2016, 133-155.

Chaudiron S., Jacquemin B., Kergosien E. « L'apport du web sémantique à la sauvegarde du patrimoine immatériel: le cas du textile, de la musique et de la mine », *Actes du 23ème Colloque Franco-Roumain en sciences de l'information et de la communication: Information, Communication et Humanités Numériques - Enjeux et défis pour un enrichissement épistémologique*, Cluj-Napoca (Roumanie), 18-20 octobre 2018, https://hal.archives-ouvertes.fr/hal-02356933.

Demonet M., Geffroy A., Gouaze J., Lafon P., Mouillaud M., Tournier M. *Des tracts en mai 1968. Mesures de vocabulaire et de contenu*, Armand Colin et Presses de la Fondation Nat. des Sc. Pol., Paris, 1975.

Favier L., « Humanities Crowdsourcing », *ZIN, Information studies*, Warsaw, Polish Librarians Association, vol. 54, n° 2, 2016, 7–21.

Fraisse A., Jenn R., Tran Q.-T., "Crowdsourcing Model for Multilingual Corpus and Knowledge Construction: The Case of Transnational Mark Twain", *ZIN, Information Studies*, vol. 56, Issue 1(111), 2018, 21-32.

Marchand P., « Quelques traces chronologiques de l'exploration textométrique », *Bulletin of Sociological Methodology / Bulletin de Méthodologie Sociologique,* 2013, *120*, 38-46.

Merzeau L., « Présence numérique: les médiations de l'identité », *Les Enjeux de l'information et de la communication*, 2009/1, 79-91.

Mustafa El Hadi W., Favier L., « Bibliothèques numériques et humanités digitales: la continuité d'une problématique sur

l'organisation des connaissances », *Les Cahiers du numérique*, vol. 11, n° 1, Paris, Lavoisier, 2015, 83-102.

Ratinaud P., Marchand P., « Des mondes lexicaux aux représentations sociales. Une première approche des thématiques dans les débats à l'Assemblée nationale (1998-2014) », *Mots. Les langages du politique, 108,* 2015, 57-77.

Smyrnaios N., Ratinaud P., « The Charlie Hebdo Attacks on Twitter: A Comparative Analysis of a Political Controversy in English and French », *Social Media & Society*, 2017, vol. 3, nº 1, [en ligne] https://doi.org/10.1177/2056305117693647

Tardy C., « Le substitut numérique: quelles mémoires muséales pour ce nouvel objet culturel? », In Idjéraoui-Ravez L. et Pélissier N. (dir), *Quand les traces communiquent… Culture, patrimoine, médiatisation de la mémoire*, Paris, L'Harmattan, 2014, 121-134.

Treleani M., *Qu'est-ce que le patrimoine numérique. Une sémiologie de la circulation des archives,* Lormont, Le Bord de l'eau, 2017.

Verlaet L., « La deuxième révolution des systèmes d'information: vers le constructivisme numérique », *Hermès La Revue*, CNRS Éditions, n° 71, 2015, 249-254.

CHAPITRE IX

Apports critiques sur les méthodologies liées aux Humanités Numériques

Christine Barats
Université de Paris, Cerlis
Julia Bonaccorsi
Université Lumière Lyon 2, UR ELICO
Amel Fraisse
Université de Lille, GERiiCO
Marta Severo
Université Paris Nanterre, Dicen-idf
Lise Verlaet
Université Paul-Valéry
Montpellier 3, LERASS

Résumé : Les sciences humaines et sociales et notamment les sciences de l'information et de la communication se sont saisies des Humanités Numériques, à la fois sous l'angle des données, des outils de traitement et de visualisation ainsi que sous celui de leur interprétation, renouvelant les questionnements épistémologiques, juridiques et éthiques. Les compétences techniques, analytiques et interprétatives développées en SIC nourrissent ainsi les apports critiques, tout en proposant des méthodologies et outils éprouvés par la démarche empirique. Si le déploiement du numérique s'accompagne d'une forte demande sociale, la posture réflexive contribue à renforcer une démarche scientifique responsable. Les enjeux et les limites des méthodologies liées aux Humanités Numériques sont ainsi régulièrement interrogés en Sic, en particulier face aux nouvelles formes de création et de dissémination des savoirs. De la collecte de « données numériques » à la production des savoirs, les méthodologies mobilisées par les chercheurs en Sic questionnent les méthodes de traitement des données. En d'autres termes, entre administration de la preuve et dimension heuristique, quels sont les enjeux et les limites des méthodologies liées au numérique ? Sans réduire ces questions à des contraintes, les apports critiques des Sic sur les relations entre techniques,

écritures, données, mémoire et savoirs sociaux, constituent autant d'axes réflexifs puissants sur les enjeux éthiques et politiques de la recherche en SHS.

Mots-clés : Approches quantitatives et qualitatives, données, traces, outils de traitement de données, outils de visualisation et d'interprétation, posture réflexive, demande sociale, recherche-action.

Abstract: Humanities and Social Studies and more specifically Communication and Information Studies (CIS) are involved in the Digital Humanities from the angle of data, processing tools and visualization as well as that of their interpretation, thereby renewing epistemological, legal and ethical questions. The technical, analytical and interpretative skills developed in CIS thus critical contributions, while putting forward methodologies and tools proven by the empirical approach. While the deployment of digital technology is accompanied by important social demand, the reflexive view makes for a stronger and responsible scientific way. The challenges and limits of the digital methodologies are therefore regularly questioned by the CIS. From data collect to knowledge production, the methodologies developped by researchers of this field analyze the data processing methods. In other words, what are the challenges and what are the limits of digital methodologies, what's the impact in the scientific practices? These questions are not only constraints, but the critical input of the CIS on the relationships between technology, writing system, data, memory, and social knowledge, offers so many potential and reflexive ways of research on the ethical and political issues of research in Social Sciences and Humanities.

Keywords: Quantitative and qualitative approaches, data, data processing tools, visualization and interpretation tools, reflective posture, social demand, action research, Digital Footprint Data.

Les Humanités Numériques face à la demande sociale

Premièrement, il s'agit d'interroger le lien entre les modalités de pratique de la recherche, notamment de financement des projets, avec le type de travaux qu'il est possible de mener et la mise en œuvre de méthodes (lien avec la demande sociale dans le cadre de recherche appliquée ; production de connaissances fondamentales).

Envisager les données numériques comme nouvelles sources (voire ressources) de savoirs implique en effet de questionner la façon dont le monde de la recherche, mais aussi celui de la demande sociale façonnent et garantissent les conditions de production de connaissances. Aujourd'hui, les pratiques de recherche en sciences humaines et sociales se trouvent de plus en plus en position d'agir dans une dynamique de recherche-action, c'est-à-dire dans une orientation stratégique ayant pour finalité l'action sociale (Reason & Bradbury, 2001), en répondant à des commandes qui proviennent d'agences de recherche à l'échelle nationale ou internationale, de collectivités territoriales, d'institutions, d'entreprises ou directement de collectifs de citoyens. De plus, l'ouverture des données et donc leur accessibilité au plus grand nombre expose les chercheurs en Sic à de nouvelles demandes sociales et à faire preuve de la même ouverture dans les méthodologies et outils employés et dans les savoirs produits. Cela est par exemple le cas des chercheurs qui appuient leurs enquêtes

sur des données issues du web et qui inscrivent leurs travaux dans le domaine des Humanités Numériques.

L'intérêt envers ces données issues du web est lié au fait qu'elles sont perçues comme des traces de l'agir social immédiatement appréhendables, sans médiation, et qu'elles peuvent donner des réponses également en temps réel, ou au moins dans des délais plus rapides que ceux des temps académiques classiques. À travers l'analyse des corpus de données issues des médias sociaux ou des échanges sur un forum ou dans un dispositif de pétitionnement en ligne, les commanditaires espèrent pouvoir connaître les comportements et les désirs des citoyens, et pouvoir mieux répondre à leurs problèmes (Severo et Romele, 2015).

Selon l'approche des sciences sociales computationnelles (Lazer *et al.*, 2009), les analyses s'appuyant sur les données numériques générées par les activités sur Internet, données considérées comme des traces du social, permettraient de construire une gestion plus performante de la société. L'étude liminaire sur la diffusion de la grippe conduite par les chercheurs de Google constitue un très bon exemple de cet « optimisme sociétal ». Ginsberg *et al.* (2009) avaient montré qu'il était possible de prévoir la diffusion du virus grippal à partir de l'étude des requêtes de Google, en construisant un indicateur plus rapide et plus fiable que les indicateurs des instituts de santé. L'analyse des traces du web apparaissait comme pouvant complètement remplacer les méthodes d'enquête classiques, plus coûteuses en termes de temps et d'argent, puisque basées sur le travail de terrain et la construction fine de l'échantillonnage des populations et des corpus étudiés.

Après une première phase d'enthousiasme inconditionné, des positions modérées et critiques se sont formulées vis-à-vis des sciences sociales dites computationnelles. Le même Lazer (Lazer *et al.*, 2014) a récemment publié une étude qui met en discussion l'efficacité de l'indicateur de Google pour la grippe et, de manière plus générale, le fait d'étudier des phénomènes sociaux à partir des données du web. Cependant, l'enthousiasme vis-à-vis de ce type de méthodes est encore très important auprès des financeurs qui voient dans les données numériques une nouvelle source d'information pour la recherche, une possibilité sans précédent de renouveler le secteur des sciences sociales et de produire des analyses plus rapides, à grande échelle et moins coûteuses. À titre d'exemple, nous pouvons citer l'appel post-attentat lancé par le CNRS en 2015 et qui a financé plusieurs études de courte durée basées sur la fouille de données du web pour analyser les traumatismes des

victimes, la propagande sur Internet, le rôle des fictions, etc. Ou encore, l'appel à manifestation d'intérêt – corpus de recherche « Grand débat national », lancé par l'ANR en mars 2019, visant à faciliter le montage de projets interdisciplinaires s'appuyant sur l'extraction d'information, la visualisation de données et l'analyse de contenus à partir du jeu des données générées par les échanges numériques entre citoyens dans le cadre du Grand débat. Plus généralement, la pratique des appels « ANR Flash » amène à réfléchir à l'affirmation d'une dynamique de recherche d'urgence qui s'appuie sur un nouveau cadre méthodologique, censé être immédiatement opérationnel.

Ces changements qui affectent la manière de faire de la recherche ne concernent pas seulement les financements spécifiquement dédiés à la recherche, ils s'observent également dans le cas de demandes d'expertises venant d'autres commanditaires publics et privés : les collectivités territoriales qui commandent des études urbaines s'appuyant sur Instagram ou d'autres médias sociaux ; les organisations touristiques qui sollicitent des études de TripAdvisor ou d'AirBnB pour connaître les habitudes des consommateurs ; ou, encore, les différents mouvements politiques s'intéressant aux échanges sur Twitter autour des élections européennes. Par ailleurs, les entreprises du web elles-mêmes, comme Facebook et Twitter, facilitent cette dynamique de recherche en ouvrant périodiquement des appels à projets pour les chercheurs intéressés par l'étude de leurs données dans un cadre de recherche-action.

Face à un tel scénario, il faut s'interroger sur le type de recherche et de méthodes mis en œuvre dans ces contextes de recherche appliquée qui s'inscrivent dans le cadre des Humanités Numériques. Serions-nous face à une nouvelle manière de faire de la recherche en Sic ? S'agit-il d'un changement épistémologique ? Et, dans ce cas, quels sont ses effets sur la manière d'étudier des phénomènes info-communicationnels ou, plus spécifiquement, des objets numériques ?

Sans aucun doute faut-il reconnaître que le changement ne touche pas simplement le type de données traitées, les techniques employées ou les outils mis en place, mais qu'il est bien plus profond. Dans le cadre de ce type d'études, en général, un processus de travail rapide s'est imposé, qui s'articule en trois phases : la collecte des données et/ou constitution de jeux de données traitables ; leur analyse, généralement à travers des méthodologies et outils dont on peut considérer qu'ils relèvent parfois de la « boîte noire » (Latour, 1987 ; Pasquale, 2015) ; enfin leur interprétation basée sur la production et

la lecture de représentations visuelles. Parfois même, cette recherche peut prendre la forme de « data sprint » ou de « hackathon », c'est-à-dire d'un travail très intensif sur une courte durée (2-5 jours) d'une équipe interdisciplinaire autour de jeux de données numériques présélectionnées (Berry *et al*., 2015 ; Munk *et al*. 2016). Sans embrasser ni l'extrême de la critique indistincte de cette démarche, ni l'extrême opposé de l'adhésion inconditionnée à certaines de ces tendances, il est fondamental de développer une réflexion critique autour de ces pratiques. En partant du constat que la recherche actuelle doit trouver des modalités de dialogue et de partenariat adaptées aux besoins et au temps de la demande sociale, il est également nécessaire de prendre du recul par rapport à une réactivité qui s'est développée sans trop être questionnée.

Les sciences de l'information et de la communication sont bien équipées pour construire ce regard critique, en portant l'attention sur trois aspects : la matérialité des dispositifs ; l'équilibre entre recherche appliquée et recherche fondamentale ; enfin, l'approche critique du numérique. Les Sic contribuent à adapter les méthodologies existantes et à explorer des nouvelles méthodes (Bonaccorsi, 2015 ; Rebillard, 2015).

Premièrement, les recherches qui s'insèrent dans le cadre décrit précédemment sont souvent basées sur un double malentendu : l'idée que les traces sont « brutes », disponibles sur Internet et celle que traces et données analysées coïncident. En effet, une approche de recherche-action qui s'appuie sur les données du web doit toujours partir du constat que les traces numériques sont construites et, par conséquent, les données qui en dérivent le sont aussi. Les dispositifs de production et de collecte de ce qui est désigné comme des traces dans la construction des questions de recherche comme dans les analyses, transforment la donnée elle-même. L'application à travers laquelle on envoie un message ou on publie une photo, ou l'outil qu'on emploie pour visualiser les données sous forme de graphe par exemple, ont un impact sur la donnée produite, modifiée et ensuite interprétée. Les Sic mettent particulièrement l'accent sur les dimensions matérielles et sémiotiques des dispositifs concernés en étudiant leurs matérialités, leurs usages et leurs contextes (Douyère, 2016) à travers des méthodologies adéquates (voir chapitre IX, en particulier 9.2).

Cela nous conduit au deuxième point d'attention pour la construction d'un regard critique vis-à-vis de la recherche-action dans le cadre des Humanités Numériques, c'est-à-dire le rapport entre recherche

appliquée et recherche fondamentale. L'opérativité qui s'est affirmée dans ces contextes de recherche invite le chercheur à privilégier l'enquête empirique et à adopter une attitude qui peut être perçue comme descriptive. Il est certain que les financements qui sollicitent des analyses du social à partir des données numériques laissent peu de place à la production de connaissances fondamentales. Généralement, ils visent à répondre à un besoin urgent de connaissance du social pour construire une action politique plus efficace. Cependant, les Sic ont déjà montré dans d'autres contextes que la recherche empirique ne peut pas être isolée et que, pour produire des résultats intéressants pour la société, elle doit s'appuyer sur un cadre épistémologique solide qui prenne en compte l'état de l'art et qui construise une théorisation à partir de la recherche appliquée (Coutant et Domenget, 2014).

Cela nous permet d'arriver au troisième aspect d'une approche de la recherche-action dans le cadre des Humanités Numériques : le fait que le cadre épistémologique soit fondé principalement sur une approche critique du numérique. Dans les sciences sociales dites computationnelles, une vision neutre et positiviste du numérique peut amener à négliger son rôle de médium. Une approche en Sic permet de porter l'accent sur la médiation des supports, des outils et des dispositifs qui entrent en jeu dans des recherches-actions qui concernent les environnements et les objets numériques.

Ainsi, la rencontre entre Humanités Numériques et recherche-action n'est pas du tout anodine. De nombreuses universités publiques et privées ont ouvert des structures, généralement appelées « media lab » ou « digital humanities lab », qui accompagnent les équipes de recherche dans ce type de démarche de recherche appliquée. Généralement, ces structures construisent une routine de recherche pour répondre aux financements publics et privés qui permet de collecter, d'analyser et de visualiser de gros corpus de données issues du web dans des délais courts et avec des coûts inférieurs à ceux de la recherche de terrain. Cependant, cette manière d'agir tend à minorer la place de la recherche fondamentale mais surtout les temporalités longues de fréquentation des mondes sociaux, des terrains et des corpus, peu compatibles avec le mode « projet ». Dans ce contexte, les savoirs des Sic sur le sens social et symbolique de la matérialité technique du numérique sont particulièrement précieux dans l'établissement de nouvelles modalités d'échange avec les commanditaires de la recherche, qui prolongent et nourrissent les questions posées de longue date par la recherche dite « appliquée », « intervention » ou « collaborative » en SHS. Selon les projets, la

définition des protocoles et des objectifs visés peut inclure une réflexion partagée avec les acteurs sociaux, les partenaires socio-économiques ou un volet dédié aux connaissances critiques et réflexives que le projet permet d'éprouver. En ce sens, plutôt que la restitution de résultats reposant sur la seule virtuosité technique et les effets de vérités produits par les dispositifs de collecte et d'analyse, il s'agit de rendre visibles les limites et contraintes, sur le mode d'un objet partagé avec la société.

Clé en main ? La mise à distance de l'instrumentation

Une seconde perspective critique concerne l'utilisation d'outils « clés en main » et l'implication dans la méthodologie d'une réflexivité sur l'instrumentation (Boullier, 2015). En d'autres termes, il s'agit de questionner les médiations techniques, comme par exemple le recours à des outils de traitement dans le cas de corpus textuels, que les outils soient existants et industrialisés ou qu'il s'agisse d'outils *ad hoc*, produits pour le besoin de la recherche. Il s'agit également de ne pas se focaliser uniquement sur les faits saillants et négliger les marges, les aberrations ou les hapax[1].

Adopter des méthodologies qui s'inscrivent dans le champ des Humanités Numériques ne se résume pas à la maîtrise de certains outils et logiciels dits « clé en main », mais implique une réflexivité sur le développement de méthodologies spécifiques, et leurs effets théoriques (Carbou, 2017 ; Bigot, 2018). Les collaborations pluridisciplinaires sont des lieux d'observation de ces mises en débat et des croisements entre des méthodes « à la main » (sans recours à des outils), et des traitements de données dont les opérations sont standardisées. Quelle délégation à la machine est-il possible d'effectuer ? Quelles formes de restitution et de mise en visibilité des données et des résultats de la recherche privilégier ? Comment articuler les approches qualitatives « à la main » et quantitatives ? Comment ne pas sous-estimer la dimension heuristique de certains outils ?

D'une part, est discutée l'adaptation des outils issus des sciences de l'ingénieur et des données aux méthodes et pratiques des sciences sociales. Ce point de vue défend que l'instrumentation choisie doit s'adapter à la situation étudiée parce qu'elle participe à la construire en tant qu'observable.

D'autre part, certains travaux visent à développer des outils logiciels spécifiques, recherchant un degré de maîtrise des effets des

1. Un hapax ne se manifeste qu'une seule fois.

programmes sur la constitution de corpus, l'archivage ou l'analyse de matériaux de recherche.

Ces outils dits « clés en main » peuvent être issus de différentes instances : des entreprises, ou agences (Linkfluence) ; des laboratoires de recherche (Iramuteq) ; des institutions patrimoniales. Par exemple, des collaborations de longue durée avec les chercheurs en Sic se sont constituées avec les institutions de savoir au service de la recherche, qu'il s'agisse à la BNF des Ateliers du web, ou de la collecte Webmédia de l'INA. Les exemples cités relèvent de trois questions de recherche, en lien avec des matériaux de granularités différentes : premièrement, il s'agit de la question des relations entre des sites web par les types de liens qui les associent ainsi que celle de leur qualification et interprétation par la visualisation de ces liens (Mazza, 2009). La deuxième question a trait à la transformation d'une unité textuelle en données. La troisième concerne l'archivage et la conservation de nouveaux types de documents. Les trois questions sont traversées par la problématique de la construction de corpus dont il s'agit de déléguer, tour à tour, la sélection et la collecte, la représentation ou la conservation à une instance médiatrice tierce.

À chaque fois, une conception du texte, du discours, du terrain d'enquête sont mobilisés et ne sont pas interchangeables, supposant de construire des parcours interprétatifs différents. Par exemple, les outils statistiques peuvent apparaître comme des méthodes d'analyse toutes indiquées pour traiter (semi)-automatiquement de grandes masses de données. Leur dimension heuristique a été maintes fois soulignée : les outils rendent ainsi visibles certains aspects des données, non observables à la main.

Certains exemples de réussite de logiciels développés (par et/ou avec des chercheurs Sic) sont à souligner, à l'instar d'Iramuteq, TextObserver, TXM ou Gephi. Ces outils posent la question de leur maîtrise, que ce soit au regard des fonctionnalités qu'ils proposent ou de leurs partis pris méthodologiques. Les outils statistiques ont ainsi montré leur dimension heuristique dans le cas de grands corpus, en particulier en Sic où l'analyse du sens de ces corpus nécessite que les textes soient situés et contextualisés.

Un enjeu important soulevé par les usages des outils clés en main dans les protocoles méthodologiques concerne les hapax, les "quantités négligeables" ou les "aberrations" écartées par certains traitements algorithmiques, lesquels ont pour objectif initial de révéler des tendances, des moyennes, des modes… Ce sont en effet des études

sur ces “aberrations”, ces “singularités”, ces “hors-normes” qui ont permis d’identifier et d’apporter des solutions à des problématiques typiquement sociales. Les SHS et les Sic font face à une variété de corpus fondamentalement polymorphes : données typographiques, iconographiques, d’images (fixes ou mobiles), de vidéos, de textes (manuscrits, transcrits, numériques), de sons, des données vectorielles, des corpus en 3D (physique : statues ; numérique : images 3D), des corpus patrimoniaux (manuscrits rares, œuvres protégées, visites virtuelles, etc.)… Cette diversité des matériaux de recherche selon les terrains n’est interopérable et traitable que par un nombre restreint de chercheurs en Sic mobilisant des méthodologies « actives » combinant des rationalités plurielles, dont celles dites computationnelles. De plus, rappelons que les tâches de collecte et de constitution des corpus sont également une étape déterminante pour les recherches en Sic et qu’elles conditionnent le traitement et l’interprétation des résultats (voir *infra* chapitre IX).

Plusieurs enjeux peuvent être distingués : concevoir et développer des outillages spécifiques, basés sur des approches computationnelles contextualisées pour et par le numérique, dans la perspective d’une cohabitation raisonnée des recherches. Autrement dit, les apports méthodologiques en Sic participent à faire évoluer les méthodologies basées sur une utilisation que l’on pourrait considérer comme « passive » d’outils clé en main, vers des méthodologies basées sur une construction « active » où les chercheurs orientent et adaptent les outils à leurs besoins. Pour cela encore, les Sic ont toute leur place. En effet, elles sont « pluri-outillées » méthodologiquement pour comprendre l’écosystème informationnel et communicationnel de ces outils et, ce faisant, trouver des réponses à une (grande) partie des problématiques citées ci-dessus.

La science ouverte ? Précautions de méthodes et possibles

Les pratiques éthiques de la recherche sont explicitées et redéfinies aujourd’hui de manière croissante depuis des cadres externes aux disciplines et aux objets traités. Ces pratiques se matérialisent dans des réglementations qui affectent les décisions en termes de méthodes. C’est notamment le cas des chartes sur l’intégrité scientifique, ou encore la réglementation européenne sur les données personnelles, la mise en place de comités d’éthiques. Plus largement, l’injonction à « la science ouverte » impacte les objectifs des projets.

Cette injonction à des formalités « open » de publicisation de la recherche en SHS est constitutive des modalités de financement par projet évoquées précédemment. Les instrumentations logicielles et leurs limites doivent être spécifiées et interprétées à l'aune de la connaissance des instances qui réalisent ces dispositifs ou encore ces architextes. Mais surtout, la critique de la recherche « numériquement équipée » concerne sa valorisation par des dispositifs médiatiques, comme des plateformes d'annotation contributives, des sites web, des carnets de recherche, voire des webdocumentaires, l'ensemble de ces exemples présentant des formats divers des données : visualisations de données ; mise à disposition de corpus ; exposition de documents numérisés (océrisation ; archives sonores). Ces différents exemples montrent une variation entre un usage interne au groupe du projet de recherche (plateforme privée) ou au contraire, largement public et intermédiatique (réseaux sociaux associés). Il s'agit une nouvelle fois de ne pas négliger l'explicitation des effets produits vis-à-vis de l'objet de recherche, et plus globalement des méthodes de recherche mobilisées.

Ces stratégies ne relèvent pas à strictement parler de la méthode : cependant, leur définition et leur formalisation, le plus souvent dans l'écriture du projet de recherche (ce que les Plans de gestion des données impliquent en partie), sont corrélées à une réflexivité nécessaire sur les statuts des données de la recherche, leurs propriétés juridiques et sociales permettant des types de circulation et diffusion plus ou moins étendus. Au-delà de ces questions, les arguments des partisans de l'ouverture des données supposent la réutilisation d'un matériau (selon une certaine rationalité économique), comme s'il était possible d'accommoder à plusieurs sauces les mêmes données (rouvrir un calcul, réinterroger les résultats d'une enquête etc.). L'ouverture des données renvoie directement à une représentation de la recherche comme un cycle, en gommant les transformations et les réécritures qui s'y opèrent. La mise à disposition des matériaux de la recherche suppose une anticipation active dans la construction et la conduite même de l'enquête et des orientations scientifiques de la transmission à autrui. L'ouverture est à questionner également à partir de la place occupée par les dispositifs de valorisation dans les projets en Humanités Numériques, comme les plateformes ou sites web, mais aussi les visualisations de données.

Conclusion

Ces dispositifs et supports qui donnent à voir et interpréter les matériaux de recherche selon des objectifs communicationnels divers

supposent une variation des qualités mêmes des objets culturels (comme traces, objets, sources) en fonction d'univers d'organisation des connaissances instituants et pluriels (patrimoine, participation citoyenne, médiation…). Les modes de valorisation des projets en Humanités Numériques sont ainsi pris dans ce paradoxe qui engage la production de technologies intellectuelles comme outils et en même temps comme médias. Par exemple, dans le cas de la numérisation de sources, la connaissance scientifique s'élabore à la croisée de l'encodage, du traitement, de l'indexation, de l'enrichissement documentaire, de l'analyse et de l'interprétation académique, mais aussi de la visibilité médiatique désirée par l'ensemble des partenaires, selon des objectifs différenciés : par exemple, l'équipe de recherche, la bibliothèque, le musée, les institutions locales, etc. Les médiations numériques des savoirs produits relèvent ainsi d'ajustements institutionnels, techniques, économiques et sémiotiques. La réflexivité scientifique doit interroger les formats médiatiques mêmes, en tant qu'ils participent de la production de la connaissance. Souhaitons que ces questions innervent les Humanités Numériques : elles concernent la représentation des *corpora* numériques en tant que savoirs dotés de nouvelles valeurs.

POSITIONNEMENT

L'instrumentation numérique de la recherche en SHS s'inscrit dans de nouvelles configurations et attentes sociales puissantes, qui engagent et contraignent les méthodes, et nécessitent une attention critique qui soit portée de manière pluridisciplinaire. Il s'agit de rendre explicite le périmètre de savoirs possibles à partir de l'analyse de données numériques collectées ou construites selon plusieurs niveaux que les Sic sont particulièrement à même d'éclairer en se saisissant des médiations info-communicationnelles : premièrement, la réflexivité méthodologique doit porter sur le postulat d'interprétation des données en traces actualisées du social, postulat qui prévaut souvent dans le cadre de projets appliqués. À un deuxième niveau, le statut des outils logiciels mobilisés doit être explicité dans la méthodologie (outils clés en main ou production *ad hoc*) ainsi que la manière dont ils interviennent dans la collecte et le traitement des données, et agissent sur les phénomènes observés. Troisièmement, l'enjeu d'une recherche ouverte et en partenariat avec des institutions extra-académiques conduit à intégrer dans la méthodologie différentes phases de partage et de diffusion des données et des analyses *via* des plateformes ou des sites web : cette circulation des savoirs en œuvre dans le temps même de la réalisation de la recherche conduit à un régime de médiatisation dont les freins, les bénéfices et les implicites communicationnels doivent être analysés.

BIBLIOGRAPHIE

Berry D. M., Borra E., Helmond A., Plantin J-C et Rettberg J. W., 2015. « The Data Sprint Approach: Exploring the Field of Digital Humanities through Amazon's Application Programming Interface ». *Digital Humanities Quarterly*. 9 (4).

Bigot J-E., 2018. *Instruments, pratiques et enjeux d'une recherche numériquement équipée en sciences humaines et sociales*. Thèse de doctorat en Sciences de l'information et de la communication, Université de Technologie de Compiègne, en ligne.

Bonaccorsi J., 2015. Quelle réflexivité médiatique pour les Humanités Numériques ?. *Les Enjeux de l'information et de la communication*, 16/2, pp. 83-97.

Boullier D., 2015. « Les trois Humanités Numériques et le positivisme algorithmique », Toulouse, 26 Nov., *Colloque Franco-québecois*. URL: https://www.canal-u.tv/video/fmsh/humanites_numeriques_dominique_boullier.20859. Dernière consultation le 12/03/2020.

Carbou G., 2017. « Analyser les textes à l'ère des Humanités Numériques. Quelques questions pour l'analyse statistique des données textuelles », *Les Cahiers du numérique*, n° 3/13, 2017, pp. 91-114.

Coutant A. et Domenget J.-C., 2014. « Un cadre épistémologique pour enquêter sur les dispositifs sociotechniques d'information et de communication ». H. Bourdeloie, D. Douyère. *Méthodes de recherche sur l'information et la communication*. Paris, Mare et Martin. https://hal.archives-ouvertes.fr/hal-01352927/document. Dernière consultation le 12/03/2020.

Douyère D., 2016. « De la recherche en communication : matérialité, critique et idéologie(s) ». *Questions de communication,* (29), pp. 195-221.

Ginsberg J., Mohebbi M. H., Patel R. S. *et al*., 2009. « Detecting influenza epidemics using search engine query data ». *Nature*, 457, 7232, pp. 1012-1014.

Latour B., 1987. *Science in action: How to follow scientists and engineers through society.* Harvard university press, 274 p.

Lazer D., Pentland A. S., Adamic L., *et al*., 2009. « Life in the network: the coming age of computational social science ». *Science* (New York, NY), 323(5915), p. 721.

Lazer D., Kennedy R., King G. *et al*., 2014. « Big data. The parable of Google Flu: traps in big data analysis ». *Science* (New York, NY) 343(6176), pp. 1203-1205.

Mazza R., 2009. *Introduction to Information Visualization*. Springer Science & Business Media, 160 p.

Munk A. K., Meunier A., Venturini T., 2016. « Data sprints: A collaborative format in digital controversy mapping » in Ribes D. and Vertesi J. (eds) *DigitalSTS: A Handbook and Fieldguide*. 458 p.

Pasquale F., 2015. *The black box society: The secret algorithms that control money and information*. Cambridge, Harvard University Press, 320 p.

Reason P. et Bradbury H., 2001. *The SAGE Handbook of Action Research. Participative Inquiry and Practice*. 1st Edition. London, Sage, 752 p.

Rebillard F. (2015). « L'étude des médias est-elle soluble dans l'informatique et la physique ? À propos du recours aux digital methods dans l'analyse de l'information en ligne ». *Questions de communication,* n°20, pp. 353-376.

Severo M., et Romele A., 2015. « *Soft data*. Essai d'une nouvelle définition des données pour les études territoriales ». *Traces numériques et Territoires, collection Territoires Numériques*, Paris, Presses des Mines. https://books.openedition.org/pressesmines/1984. Dernière consultation le 12/03/2020.

CHAPITRE X

Computation et interprétation
Les études digitales une opportunité pour penser une rencontre avec les Sic

Franck Cormerais
Université Bordeaux-Montaigne, Mica

Résumé : Le terme computation renvoie au calcul (*computus*), à la mesure et à la précision, tandis que l'interprétation favorise l'idée d'écart, de discussion, de débat. L'interprétation exprime une « prestation intermédiaire » (*inter-pretatio*), où le sens ne peut être donné ou actualisé sans elle. Disons qu'avec les Humanités digitales la tension entre computation et interprétation dans le cadre des Sic se pense d'une double manière. D'abord comme un dépassement d'une opposition entre une herméneutique du soupçon et une herméneutique de la confiance (Ricœur, 1969). Ensuite, comme un déplacement des limites de l'interprétation. Dans un environnement digital, la relation entre calcul et interprétation change. En effet, les conditions de l'interprétation se transforment dans la relation qui s'établit entre une herméneutique du digital d'un côté, héritière d'une « sémiosis hermétique » (Eco, 1992) et d'une science quantitative des données de l'autre.

Mots-clés : études digitales, interprétation, raison computationnelle.

Abstract : The term computation refers to computus, measurement and precision. The term interpretation implies the idea of gap, discussion, debate. The interpretation expresses an "intermediate benefit" (inter-pretatio), where meaning cannot be given or updated without it. Let's say that with the digital humanities the tension between computation and interpretation in the context of Sic is thought of in a double way. At first as an overcoming of an opposition between a hermeneutic of suspicion and a hermeneutic of trust (Ricœur, 1969). Then, as a shift of interpretation boundaries. In a digital environment, the relationship between calculation and interpretation changes. Indeed, the conditions of interpretation are transformed in the relationship that is established between a digital hermeneutic on one side, heir to an «hermetic semiosis» (Eco, 1992) and a quantitative science of the data on the other.

Keywords: digital studies, interpretation, computational reason.

Dans le milieu techno-sémiotique du digital, le travail de la pensée s'organise autour d'une relation entre *mathêmata* (au double sens de connaissance et de formule de la connaissance) qui s'incarne aujourd'hui dans les algorithmes, et *deixis* (au sens grec d'action de montrer et plus largement de communiquer). Un ensemble d'opérations, à la croisée de la computation et de l'interprétation rebat alors les cartes entre approche sémantique et approche critique, afin d'éviter aussi bien la disparition des faits que la « surinterprétation ». Nous envisagerons le déploiement dans les Sic des études digitales autour de la question de la manipulation des langages formels (Partie 1), de l'apparition d'un régime de l'hyperdocument (Partie 2), mais aussi d'une convergence de l'info-doc et de la communication sociale (Partie 3).

La computation et les algorithmes, ce qui se donne à interpréter

L'enjeu de la rencontre du texte et plus généralement du document avec les données informatiques se joue dans une *critique de la raison computationnelle* (Bachimont, 2008). Il s'agit bien d'aller au-delà de l'informatique conçue comme l'automatisation des raisonnements menés dans des systèmes algébriques formalisant le raisonnement logique, dont l'apparition moderne remonte à Leibniz. Ce qui est généralement reproché à l'informatique est une manipulation aveugle de signes considérés indépendamment de leur contenu permettant de mener à bien les calculs. Dans ce cas, un programme informatique met en œuvre une manipulation de règles formelles indépendamment du contenu signifié. Le résultat du programme devient intelligible et précis car à chaque signe du symbolisme informatique est attachée une entité du monde mathématique. Le primat d'une signifiance écarte ainsi le signifié.

Comme il est difficile de se satisfaire d'une telle approche, une deuxième façon d'envisager l'informatique s'impose alors. Il ne s'agit plus de manipuler des symboles associés à un modèle mathématique,

mais des symboles librement interprétés selon des règles empruntées à la langue générale ou spécialisée d'un domaine d'activité. Cette perspective s'ouvre à la compréhension de « ce que nous faisons avec nos machines informatiques quand il ne s'agit pas seulement d'automatiser nos symbolismes algébriques issus des sciences mathématiques, mais de mécaniser la manipulation de symboles comme s'ils étaient algébriques (c'est-à-dire en congédiant leur signification pendant leur manipulation) alors qu'ils ne le sont pas (les comprendre, c'est les considérer dans leur matérialité physique avec leur signification sans jamais les dissocier) » (Bachimont, 2008, p. 4).

Ce double mouvement situe les Sic et les études digitales dans la convergence d'une numératie (calculabilité) et d'une littératie (information écrite et transcrite numériquement) où aux combinatoires algébriques, « il faut ajouter l'invention que permettraient ces combinatoires pseudo-algébriques. Pour aborder cette invention, nous proposons de considérer ces formalismes comme des écritures et des outils intellectuels, c'est-à-dire comme des symboles dont la manipulation peut être purement formelle si on en reste à leur nature matérielle et physique, mais dont l'interprétation ne peut pas se rapporter à des règles algébriques » (Bachimont, 2008, p. 5).

Dans ce contexte, les Sic retrouvent à travers les études digitales l'histoire des outils intellectuels (Goody, Ong, Otlet, Estivals, Escarpit, etc.). Ces outils, en se mettant au service du sens, s'inscrivent dans un mouvement d'extériorisation et de prothétisation de la pensée. Si l'écriture autorise la synthèse du temps dans l'espace, en permettant d'inscrire ce qui est dispersé dans le temps (la parole), l'informatique permet le déploiement de l'espace en temps. En effet, un programme n'est pas autre chose qu'un dispositif réglant un déroulement dans le temps, par le calcul ou l'exécution, à partir d'une structure spécifiée dans l'espace : l'algorithme.

Celui-ci est donc un moyen de « certifier » l'avenir, d'en éliminer l'incertitude et l'improbable pour le rapporter à la maîtrise. « Le temps de l'informatique n'est donc pas une disponibilité à ce qui va venir, aussi improbable que cela puisse être, mais la négation du futur dans son ouverture pour le réduire à ce qui peut s'obtenir à partir du présent » (Bachimont, 2008, p. 11).

L'algorithme propose une équivalence entre temps et espace : le temps devient synonyme de l'exploration systématique d'un espace de calcul. Comme parcours de tous les cas possibles, l'algorithme produit un espace parcouru en un certain nombre d'étapes, spécifiées par le

calcul. Le problème posé par la computation est bien celui des limites de la compréhension des faits et du renouveau de l'interprétation. La raison computationnelle appelle donc une nouvelle critique pour penser les outils intellectuels.

Si le calcul possède pour vertu de réduire la complexité en permettant d'explorer systématiquement un espace dans lequel se trouve le texte ou le document écrit, il faut noter cependant que le problème du calcul et de l'intelligibilité de ses résultats demeure un point aveugle. En effet on se trouve confronté, selon des protocoles et des calculs compliqués, à des connaissances obtenues par l'interaction de sources hétérogènes, parfois incomplètes ou encore fausses. Ignorant quelquefois même les sources mobilisées, ne contrôlant pas les calculs menés : le sujet connaissant se trouve devant une masse d'informations et de documents qui ne sont pas a priori intelligibles. Cette situation a d'ailleurs tendance à se généraliser au concepteur de programme, par exemple, dans le cas d'un apprentissage profond réalisé par ce programme (*Deep Learning*).

Nous sommes en droit alors de demander : « Comment pouvons-nous nous les rendre intelligibles ? En les rapportant aux concepts et structures rationnelles que nous connaissons et maîtrisons déjà. Autrement dit, c'est en rapportant les résultats calculés à nos connaissances mémorisées que ces résultats seront appropriables et intelligibles. C'est donc à une articulation entre calcul et mémoire qu'il faut arriver » (Bachimont, 2008, p. 13).

Pour éviter la situation paradoxale d'un déficit d'intellection avec la croissance de la puissance du calcul, il convient d'articuler les résultats calculés en les rapportant à nos connaissances mémorisées lors du passage du document à l'hyperdocument. En conférant aux inscriptions une dynamicité propre, l'hyperdocument, par son passage par la computation, met en évidence une forme symbolique originale. Il autorise une expérience intellective par manipulation dans une extériorité qui s'expérimente dans des inscriptions. Ce que donne à penser la rencontre de l'hyperdocument et du calcul est la limite qu'il faut poser aux conditions d'une surinterprétation par une algorithmisation qui ferait fi des traditions de l'herméneutique et de la rhétorique. Une relégation du sens dans un appareil signifiant (*computer*) consisterait alors à « algorithmiser » la lettre, c'est-à-dire à l'empêcher de fonctionner comme signe.

L'hyperdocument ou le croisement de la lettre (interprétation) et du calcul (computation)

Avec l'hyperdocument, les Sic peuvent aborder le processus interprétatif comme dimension symbolique, où se croisent langue technique du calcul et langue de tradition autrefois opposées (Heidegger, 1990). Avec l'hyperdocument, il s'agit de revenir sur l'utilisation des formalismes dans des contextes où la signification des symboles manipulés n'est pas rigoureusement déterminée. Il faut remonter aux années 90 pour qu'apparaisse le mot hyperdocument (Balpe, 1990). L'hyperdocument dérive de l'hypertexte dans le projet Xanadu (Ted Nelson 1965). L'hyperdocument réalise la synthèse suivante :

> il désigne indifféremment un hypertexte ou un hypermédia, c'est-à-dire un produit, à créer ou à consulter, dont le contenu est informatisé (toutes les informations sont numérisées) et organisé de manière non linéaire. L'utilisateur d'un hyperdocument effectue un parcours interactif à travers cet ensemble d'informations textuelles, sonores et iconographiques. En passant d'une information à l'autre, au gré de ses envies et de ses besoins, il constitue un réseau de liens entre divers fragments de savoirs et construit sa propre connaissance. (Godinet, 1995, p. 91).

Il paraît utile de proposer une première approche d'une généalogie qui souligne l'évolution vers l'hyperdocument (tableau n° 1). L'hyperdocument apparaît alors comme une hybridation du document et du semi-document ; il superpose des caractéristiques qui mobilisent des objets, des techniques, des pratiques, des espaces dédiés, etc.

Tableau 1 – Éléments pour une généalogie de l'hyperdocument

Caractéristiques	Document	Semi-document	Hyperdocument
Objets	Livre, manuscrits, revues, journaux	Musique, radio, film, Télévision	Fichier numérisé, Hypertexte
Supports	Papier	Disque, CD, DVD	Écran, serveur
Éditions	Imprimerie	Pressage	Mise en ligne
Modalités	Écriture, lecture	Enregistrement	Donnée structurée
Consultation	Disposer	Visionner, écouter	Interroger
Lieu	Bibliothèque	Médiathèque	Ubiquithèque

Ce tableau appelle l'écriture d'une généalogie plus complète qui viendrait enrichir celle de l'audiovisuel (Perriault, 1981) et celle de l'informatique (Lazard & Mounier Kuhn, 2016). Les Sic trouveraient dans une telle opération un champ empirique pour déployer une archéologie des médias (Parikka, 2018). Les hyperdocuments traduisent de profondes mutations dans les conditions de production et de réception de la connaissance. Ces dernières s'appréhendent avec le concept de « document numérique » (Cotte, 2004), qui précise l'incidence des techniques, en lien avec le support, et annonce le renouvellement des méthodes d'analyse et d'interprétation (Cormerais & alii, 2016). Dans la phase actuelle, distinguons trois régimes de l'hyperdocument (Tableau n° 2) :

Tableau 2 – Petite généalogie de l'hyperdocument

Régime documentaire	Réseau	Technologie informationnelle	Opération
Hyperdocument (Hyper 1)	Web 1.0	Moteur de recherche	Indexation
Hyperdocumentalité (Hyper 2)	Web 2.0	Réseaux Sociaux	Traçabilité
Hyperdocumentarité (Hyper 3)	Web 0.3	Algorithmes, internet des objets	Filtrage des données massives

L'hyperdocument (Hyper 1) bouleverse l'accès du contenu du document ; il modifie les conditions de la consultation. L'hyperdocumentalité (Hyper 2) souligne l'importance du document comme objet d'une construction sociale (Ferraris, 2016). Avec l'hyperdocumentarité (Hyper 3) se pose directement la question de la relation entre le document et la donnée. Un mouvement d'effacement du document qui n'est qu'apparent s'opère avec les données massives (*Big data*). Dans ce contexte, l'hyperdocument appelle un *principe de documentarité* (Crozat, 2016) qui implique une réflexivité posée sur le document à partir d'un triple point de vue : a) culturel (transmission, preuve), b) technologique (disponibilité), c) scénaristique (lisibilité). Cette unité triplice nous ramène à de nouvelles conditions d'interprétation qui rassemblent langage informatique et sémiotique.

Avec la documentarité « le but n'est donc pas de dire ce qui est mieux ou souhaitable, mais de se poser en situation réflexive sur nos pratiques de lecture, sur nos pratiques documentaires en général, de s'interroger sur le statut des contenus que l'on manipule et que l'on produit » (Crozat, 2016). L'hyperdocumentarité s'interroge sur les pratiques de

lecture et apporte un complément à l'approche fonctionnelle et non substantielle du document (Briet 1951, Otlet 1934, Buckland, 2017). L'intérêt de l'hyperdocumentarité est donc une posture qui permet d'évaluer ce qui fait document dans un objet informationnel. Ce déplacement de l'hyperdocumentarité vers la réception est important car, comme le souligne Bachimont, « le document, du point de vue de la production, est dissout en données, dans les deux sens : linked data comme informations élémentaires, ou ressources annotées, et big data comme éléments infra-informationnels qui ne prennent un sens d'information que dans la masse à laquelle ils contribuent et qui seule porte des informations. Du coup, c'est bien au niveau de la réception que le sens se construit » (Bachimont, 2016).

C'est précisément du côté de la réception, de la *lettrure* (Souchier, 2012), autre nom de la littératie, qu'il faut aller pour comprendre le doublet computation/interprétation du document. Nous aurons alors une meilleure compréhension du redoublement du document numérique par l'hyperdocumentarité.

Le déplacement des limites de l'interprétation : l'hyperdocument dans ses milieux de concrétisation et de communication

L'hyperdocumentarité représente un moment dans lequel les Sic trouvent leur place dans le champ des sciences. Il s'agit, après des pionniers tels Robert Escarpit (1976) et Robert Estivals (1993), d'utiliser des modes d'étude en s'appuyant sur des formes de connaissances qui allient la computation et l'interprétation.

Afin de mieux prendre en compte un nouveau conflit des interprétations (après le « tournant linguistique »), entre le chiffre et la lettre (numératie *versus* littératie), les Sic doivent envisager le redoublement d'une lecture projective (interprétative) et d'une lecture objective (computationnelle) des hyperdocuments. La lecture projective est liée au statut du sujet, tandis que la lecture objective repose sur la précision apportée par l'ordinateur. Face aux risques d'une réduction du langage de tradition (herméneutique, rhétorique, poétique) par les données autrement dit par le primat du calcul, la rencontre du signe et de la donnée prend en charge l'adressage du signifié par un code relié à la puissance du calcul.

Pour aller vers une conception élargie des « études digitales » (une revue porte d'ailleurs ce nom), il faut ouvrir le document à un processus communicationnel qui excède l'approche ontologique

de l'informatique comprise comme un ensemble structuré par des métadonnées. Pour cette raison, il paraît souhaitable d'aborder l'hyperdocument dans une inscription interdisciplinaire et de renouveler l'aventure d'une écriture collective menée autrefois par le collectif Pédauque (Sultan & Salaün, 2010). Pour renforcer cette orientation, il convient de conditionner la relation entre le signe et la donnée à la réalisation d'une synthèse des trois formes majeures du donné (le conscientiel, l'informationnel, l'automatisé). Nous voyons là une possibilité de produire un mixte d'activité mentale (symbolique) et d'activité machinique (calculatoire), de littérarité (importance de la lettre) et de numérarité (importance du calcul). Pour réaliser ce programme, il faut réunir au moins deux conditions :

Tableau 3 – Conditions d'une approche des Hyperdocuments

Première condition	Effet
Ouvrir la rationalité descriptive du document à une logique de l'usage socialisé des inscriptions par et dans les disciplines	Au-delà des ontologies, l'analyse débouche sur un croisement du logiciel et des modes d'existence disciplinaires
Seconde condition	**Effet**
Relier le document à une pluralité d'interprétation pour organiser une nouvelle exégèse	Établir une herméneutique matérielle d'une technoculture partagée

La valeur de l'hyperdocumentarité induit un enrichissement du contexte du sens par la technique et produit une nouvelle éditorialité. Pour atteindre une reconfiguration technoculturelle de la connaissance avec les Sic, le contexte (*contextus*) s'envisage comme un milieu de l'expression et de l'action, où la fonction documentaire reconfigure une connaissance reposant sur l'invention d'une herméneutique critique du texte calculé.

Conclusion : le retour de l'information dans les Sic au-delà des partages anciens

Avec la documentarité s'ouvre un processus informationnel et communicationnel important pour les Sic, nouvelle conditionnalité de production du sens par la prise en compte des appareils contre une conception éthérée d'une science humaine sans instrument. La position interdisciplinaire des Sic favorise le questionnement des grands partages et trouve là une possibilité de réconcilier la division entre information et communication (Bougnoux, 1995). Dans ce

processus, il n'y a pas d'antinomie entre une approche statistique et une approche sémantique par adressage des données. L'invention d'un langage écriture sert d'interface entre les langues naturelles et formelles pour utiliser les possibilités computationnelles de la manipulation des symboles.

La question de l'interprétation, dans le milieu technosémiotique du digital, est soumise à une révision des postulats de l'herméneutique et à un renouveau de la pensée critique ; la relation entre l'ordre des lettres et des chiffres appelle, par le biais du calcul, un dépassement du grand partage entre les sciences humaines et les « sciences du numérique »[1] (connue sous l'abréviation ISN dans l'enseignement).

Ces sciences possèdent un contour flou et même une allure technophilique. Le programme officiel de ces sciences, révisé en 2017-2018 dans les lycées, est construit autour de quatre parties : représentation de l'information, algorithmique, langages et programmation, architectures matérielles. Cette orientation laisse une place congrue à la culture des textes et aux sciences humaines et sociales. À l'inverse de cette démarche, les Digital Humanities valorisent une approche pluri- et transdisciplinaire dans les Sic, où l'accès à l'information scientifique et technique interroge conjointement la transformation des supports et les modes de production des connaissances. C'est la question de l'articulation entre l'interprétation et la computation.

L'exploitation de nouvelles potentialités « herméneutiques » du digital (notamment le développement des standards d'annotation et d'éditorialisation) invite à être examinée au prisme d'une « logicialisation » et d'une « délégation logicielle ». Cette herméneutique soulève de nombreuses interrogations dans l'horizon d'une pensée critique.

La première interrogation porte sur le risque d'un écrasement de l'appréhension du sens en contexte par le calcul probabiliste. Devant une rationalité computationnelle, une mise en garde contre l'excès de confiance dans la force objectivante des données ou contre l'hégémonie d'une conception quantitative de la preuve s'avère nécessaire. Les médiations numériques dans leur matérialité appellent donc une analyse multidimensionnelle qualitative et quantitative.

1. L'introduction en 2012 dans les lycées des sciences du numérique introduit une séparation entre les littéraires et les scientifiques. En effet, les sciences du numérique rejoignent les enseignements de spécialité suivants : *mathématiques*, *sciences physiques et chimiques*, *sciences de l'ingénieur* et *sciences de la vie et de la Terre*.

Une nouvelle interrogation se pose face à l'industrialisation de la logistique scientifique. En effet, les médiations sont organisées par un appareillage qui combine des intérêts, des capacités d'agir qui résultent toujours d'actions dirigées vers une conception de la connaissance et des enjeux de pouvoir. Au conflit des interprétations se superpose un conflit économique entre plusieurs catégories d'acteurs : les opérateurs de télécommunication, les sociétés d'ingénierie logicielle et les producteurs de matériel informatique. Ces acteurs ont un intérêt évident, et aussi différent de celui du monde de la recherche académique, qui suppose des instruments de plus en plus sophistiqués.

La connaissance, enfin, n'échappe pas à l'interrogation relative à une tendance certaine à la marchandisation. Les risques liés aux formats sont patents sur les plateformes d'intermédiation. C'est pourquoi une réflexion sur les modèles économiques et les stratégies éditoriales, en relation avec l'évolution du statut d'auteur et des droits de propriété, sont une priorité pour empêcher la formation de rentes. L'opposition entre logique « propriétaire » et « libre diffusion » des savoirs scientifiques (*open data*) ne conduit pas à une entière disparition des finalités économiques, mais plutôt à une réflexion sur le statut de la connaissance comme bien particulier, sinon comme bien commun (Pélissier, 2017). Pour cela, l'attention portée à l'action des infrastructures et des dispositifs du milieu digital nécessite une perspective en termes d'économie politique de la production scientifique.

Notre réflexion autour de l'hyperdocument nous a conduit à comprendre l'évolution des infrastructures réticulaires et des opérations réalisées par les technologies informationnelles associées. Avec l'hyperdocument, le franchissement d'une étape dans nos rapports aux objets techniques se réalise dans la médiation des connaissances et leur production. Avec l'hyperdocument apparaît un régime documentaire ne reproduisant pas le grand partage historique, interne aux Sic, entre information et communication.

L'hyperdocument rend possible une hyperdocumentarité qui renvoie à des transformations de la structuration de l'information et à la (re)documentarisation. Si le document se comprend comme une « spécialité sociale de l'écriture » (Ferraris, 2016, p. 51), l'hyperdocumentarité s'entend comme une capacité réflexive et critique à produire des hyperdocuments en ajoutant aux dimensions culturelles et technologiques, une dimension économique et politique. Ainsi, nous pouvons associer au document une « économie

scripturaire » dont rendent comptent des opérations techniques liées à l'inscription et l'excription de la signifiance. Ces opérations débouchent sur un processus de la valorisation et pour cela, les Sic associent à l'approche heuristique du chercheur une approche objective reposant sur la computation. Pour réaliser ce programme, une conception élargie des études digitales s'avère nécessaire en ouvrant l'hyperdocument à un processus communicationnel qui excède l'approche ontologique de l'informatique structurée par des métadonnées. Cette orientation revient sur la relation entre le signe et la donnée pour réaliser une synthèse des trois formes majeures du donné : le conscientiel, l'informationnel, l'automatisé.

POSITIONNEMENT

L'interprétation dans le milieu technosémiotique du digital est soumise à une révision des postulats de l'herméneutique et à un renouveau de la pensée critique ; la relation entre l'ordre des lettres et des chiffres appelle, par le biais du calcul, un dépassement du grand partage entre les sciences humaines et les « sciences du numérique » (connues sous l'abréviation ISN dans l'enseignement), lesquelles possèdent un contour flou et même une allure technophilique.

Le programme officiel de ces sciences, révisé en 2017-2018 dans les lycées, est construit autour de quatre parties : représentation de l'information, algorithmique, langages et programmation, architectures matérielles. Cette orientation laisse une place congrue à la culture des textes et aux sciences humaines et sociales.

À l'inverse de cette démarche, les Digital Humanities valorisent une approche pluri- et transdisciplinaire, où l'accès à l'information scientifique et technique interroge conjointement la transformation des supports, les modes de production des connaissances et leurs implications économiques.

BIBLIOGRAPHIE

Bachimont B., 2008. « Signes formels et computation numérique : entre intuition et formalisme, Critique de la raison computationnelle », (16 pages), traduction française de « Formal Signs and Numerical Computation : Between Intuitionism and Formalism. Critique of Computational Reason » dans H. Schramm, L. Schwarte, & J. Lazardzig (Eds.), *Theatrum Scientiarum* : *Instruments in Art and Science, on the Architectonics of Cultural Boundaries in the 17th Century*. Berlin, Walter de Gruyter Verlag, pp. 362-382. www.utc.fr/~bachimon/Publications_attachments/Bachimont.pdf. Dernière consultation le 12/03/2020.

Bachimont B., 2016. https://aswemay.fr/co/010013.html, commentaire à Stéphane Crozat.

Balpe J.-P., 1990. *Hyperdocuments*. Paris, éd. Eyrolles, 216 p.

Bougnoux D., 1995. *La communication contre l'information*. Paris, Hachette, 144 p.

Briet S., 1951. *Qu'est-ce que la documentation ?*, Édit, Paris, 48 p.

Buckland M. K., 2017. *Information and society*. Cambridge, Massachusetts, The MIT Press. The MIT Press essential knowledge series, 232 p.

Cormerais F. , 2016.« Traitement des textes, sens et logique des formats », *Études digitales* n°1, 2016, Classiques Garnier, Paris, pp. 25-40.

Cormerais F. Le Deuff O., Lakel A. et Pucheu D., 2016. « Les Sic à l'épreuve du digital et des Humanités : des origines, des concepts, des méthodes et des outils », *Revue française des sciences de l'information et de la communication* [En ligne], 8 | 2016, mis en ligne le 23 mars 2016, consulté le 31 mai 2019. URL : http://journals.openedition.org/rfsic/1820 ; DOI : 10.4000/rfsic.1820. Dernière consultation le 12/03/2020.

Cotte D., 2004. « Le concept de « document numérique », *Communication & Langages*, n° 140, pp. 31-41.

Crozat S., 2016. https://aswemay.fr/co/010013.html.

Eco U., 1992. *Les Limites de l'interprétation*, Paris, Grasset, 411 p.

Escarpit R., 1976. *Théorie générale de l'information et de la communication*. Paris, Hachetten 220 p.

Estivals R., 1993. Encyclopédie d*es Sciences de l'*écrit, Paris, Retz, 576 p.

Ferraris M., 2012. *Documentality: Why It Is Necessary to Leave Traces*. Oxford, Oxford University Press, 392 p.

Godinet H., 1995. « Hypertexte, hypermédia, hyperdocument… dans les activités de lecture- écriture », *Revue de l'EPI* (*Enseignement Public et Informatique), EPI*, pp. 89-100 https://edutice.archives-ouvertes.fr/file/index/docid/43551/filename/b77p089.pdf. Dernière consultation le 12/03/2020.

Heidegger M., 1990. *Langue de tradition et langue technique*, Bruxelles, Lebeer-Hossmann, 56 p.

Lazard E. et Mounier-Kuhn P., 2016. *Histoire illustrée de l'informatique*. Les Ulis, EDP Sciences éd, 304 p.

Otlet P., 2015. *Traité de Documentation : le livre sur le livre*. Paris, Mundaneum, 1934, Réédition en fac-similé, Bruxelles, Les Impressions Nouvelles, 577 p.

Parikka J., 2018. *Qu'est-ce que l'archéologie des médias ?* Grenoble, UGA Éditions, 324 p.

Pélissier-Thériot M. « Les communs culturels : nature, enjeux et perspectives socio-économique », in *Art, culture et communication : métamorphoses numériques* (dir. Pélissier et Auroy, CNRS éditions, Paris, pp. 29-40.

Perriault J., 1981. *Mémoires de l'ombre et du son. Une archéologie de l'audiovisuel*, Paris, Flammarion, 288 p.

Ricoeur P., 1969. *Le Conflit des interprétations, essais d'herméneutique*. Paris, Seuil, 500 p.

Souchier E., 2012. La « lettrure » à l'écran, Lire & écrire au regard des médias informatisés, *Communication & langages*, 2012/4 (n° 174), pp. 85-108.

Sultan F. et Salaün J-M., 2010. « Roger T. Pédauque, l'aventure d'une écriture collective », http://base.d-p-h.info/fr/fiches/dph/fiche-dph-8220.html. Dernière consultation le 12/03/2020.

PARTIE IV

Formation aux Humanités Numériques : quelle place et enjeux pour les Sic ?

INTRODUCTION PARTIE IV

Formation aux Humanités Numériques
Quelle place et enjeux pour les Sic ?

Dans cette partie nous discutons l'apport et la place des Sic sous l'angle de la formation, de l'éducation et de la sensibilisation des publics aux environnements et Humanités Numériques. Les réflexions développées convoquent l'ensemble des acteurs et publics impliqués par ce domaine émergent : les élèves de l'enseignement primaire et secondaire, les étudiants, les salariés et autres adultes en formation continue (notamment ceux qui sont concernés par des problématiques communicationnelles), mais aussi les formateurs, enseignants et enseignants- chercheurs universitaires. Il s'agit également d'identifier et de réfléchir sur l'ensemble des formations en Humanités Numériques au niveau national.

Dans le **chapitre XI**, Sylvie Alemanno et Elise Maas questionnent prioritairement les *Humanités Numériques au Travail* (H.N.T.), leurs usages et les éléments qui les caractérisent. La digitalisation d'une organisation ne se constitue pas ou peu par des pratiques intuitives, mais par un socle de compétences structurées et en adéquation avec les attentes de ses collaborateurs. De ce fait, les diplômés en Sic sont perçus par les employeurs comme devant maîtriser efficacement ces compétences, notamment au plan de la création, du développement et du management des supports communicationnels. Ce constat implique la recherche d'une adéquation des compétences info-communicationnelles requises chez les diplômés avec les attentes de leurs futurs employeurs concernant la transformation numérique de leur organisation. Les auteures suggèrent donc que la compréhension de

la complexité des Humanités Numériques au travail ne peut s'effectuer qu'à l'aune de la prise en compte de l'évolution des usages privés et sociaux, lesquels ont des incidences certaines sur les pratiques numériques en milieu professionnel. Cet éclairage permettra aux formations et diplômés en Sic de mieux appréhender les enjeux des Humanités Numériques au travail.

Le **chapitre XII** qui fait suite présente quant à lui un panoptique des formations aux Humanités Numériques au plan national, ainsi qu'une tentative de modélisation de ces cursus. Orélie Desfriches et Elie Allouche identifient les critères généraux de ces formations, puis proposent plusieurs types de visualisation des données recueillies par le biais de cartographies issues de la base Dariah et du service en ligne ministériel « Trouver mon master ». Ces bases ont été questionnées par type de formation, par discipline, par objets et par techniques. Enfin, ils exposent un modèle des principales dimensions dispensées au sein de ces formations à partir d'un panel diversifié de cursus en Humanités Numériques à l'international. Comme le soulignent les auteurs, leurs réflexions à portée exploratoire appellent des enquêtes complémentaires qualitatives auprès d'enseignants-chercheurs, ainsi qu'une étude comparative de ces formations au plan international. Toutefois le modèle présenté peut servir de base solide d'échanges entre enseignants et chercheurs intéressés par ces questions.

Dans le **chapitre XIII** qui clôt cette partie, Emilie Remond, Laurent Collet et Thibaud Hulin réfléchissent sur le positionnement, les objectifs et rôle des formateurs aux Humanités Numériques. Lorsque l'on observe l'évolution du secteur de la formation, celui-ci se conforme de plus en plus nettement à une logique de marché qui sous-tend des pratiques concurrentielles traduisant une vision utilitariste du savoir devant satisfaire en priorité les objectifs des acteurs économiques dominants. Les auteurs proposent un panorama de l'ensemble de l'offre de programmes de formation de formateurs aux Humanités Numériques, puis développent une réflexion sur la fonction de ces formateurs dans leur contexte professionnel. Il apparaît alors que les compétences associées à leur travail ne se limitent pas à la maîtrise d'aspects techniques, mais impliquent aussi la transmission d'une *culture numérique* renforçant le libre arbitre, la responsabilité des citoyens, ainsi qu'un esprit créatif leur permettant de s'adapter en continu aux évolutions socioéconomiques de sociétés hyperconnectées, évolutions de moins en moins prévisibles par les acteurs concernés.

En guise de **conclusion** de cet ouvrage mais aussi d'ouverture à d'ultérieurs échanges, Arnaud Mercier, Nicolas Pélissier et Éric Triquet traitent de la nécessité de mettre en place un socle commun de compétences et connaissances au sein des différents cursus formant aux écosystèmes numériques. Car loin de la vision irénique d'un Web vecteur de liberté individuelle et de progrès social, ce dernier a engendré plusieurs menaces affectant le développement de l'espace démocratique. Pour ces auteurs, il s'agit essentiellement des *fake news*, de l'impact négatif du numérique au plan des usages professionnels ainsi que de l'utilisation peu contrôlée et excessive des données personnelles. Pour limiter voire éviter ces dérives, il s'agira de mettre la notion de *citoyenneté numérique* au centre de toute formation traitant de cette dimension, formation devant reposer sur quatre piliers centraux : l'histoire et les enjeux du numérique ; les principes et techniques des écritures et de l'édition numérique ; les bases du codage et de l'algorithmie ; enfin, les techniques de veille, de traitement et d'interprétation des données. Dans ce contexte, les enseignants-chercheurs en Sic occupent de facto une place privilégiée pour initier, développer et dispenser des formations originales basées sur ces principes élémentaires.

Franck Debos et Nicolas Pélissier

CHAPITRE XI

Humanités Numériques au travail et professionnalisation des communicateurs

Sylvie Alemanno

Conservatoire National des Arts et Métiers (CNAM), Laboratoire DICEN IDF, France

Élise Le Moing-Maas

Institut des Hautes Études des Communications Sociales (IHECS), laboratoire Protagoras (IHECS), Belgique, laboratoire PREFics, université Rennes 2, France

Résumé : dans ce chapitre, nous évoquerons la professionnalisation des communicateurs au regard du contexte des transformations numériques qui irriguent l'organisation à tous les niveaux. Nous questionnerons l'intégration et la compréhension des outils et dispositifs numériques dans les formations en Sic, tant au niveau du développement de compétences des communicateurs que de leur compréhension des enjeux communicationnels des Humanités Numériques au Travail (HNT) telles que nous les définirons dans ce chapitre. Quels ajustements entre savoir-faire techniques, compétences stratégiques et savoir-être liés aux Humanités Numériques en environnement de travail mais également influencés par les pratiques privées et sociales ? Nous mettrons en perspective les résultats des travaux des chercheurs en Sic, et plus particulièrement ceux du groupe de recherche Org & Co, avec l'analyse de baromètres et d'études (EUPRERA, EACD, UDA, Entreprise et Média) consacrés à l'évolution des métiers de la communication et des compétences numériques nécessaires aux communicateurs. Nous pourrons ainsi formuler quelques propositions propres aux Sic sur notre positionnement quant à la professionnalisation des communicateurs induite par les Humanités Numériques au travail.

Mots-clés : Professionnalisation, Humanités Numériques au Travail (HNT), formation, communicateur.

Abstract : In this chapter, we will discuss the professionalization of communicators in the context of the digital transformations that permeate the organization at all levels. We will question the integration and understanding of digital tools and

devices in Communication and information sciences training both in terms of skills development for communicators and their understanding of the communication issues of Digital Humanities at Work (DHW) as we will define them in this chapter. What are the adjustments between technical know-how, strategic skills and soft skills related to digital humanities in the work environment but also influenced by private and social practices? We will put into perspective the results of the work of Communication and information sciences researchers and more particularly those of the Org&Co research group concerning DHW. We analyse barometers and studies (EUPRERA, EACD, UDA, Entreprises et Média) devoted to the evolution of communication professions and in particular to the digital skills required by communicators. We will thus be able to formulate some proposals specific to the Communication and information sciences on our positioning regarding the professionalization of communicators induced by digital humanities at work.

Keywords: professionalization of communicators, training, digital humanities at work (DHW).

Dans le domaine des sciences humaines et sociales, les enjeux de la formation universitaire se confondent avec ceux de la professionnalisation à tous les niveaux du L.M.D.[1], et les réflexions que suscite le sujet s'étendent déjà sur une vingtaine d'années. Dans la discipline des Sciences de l'Information et de la Communication (Sic), la professionnalisation des communicateurs est au cœur des formations, afin d'assurer une employabilité des diplômés universitaires pour répondre aux besoins d'organisations variées couvrant toutes les orientations (nombreuses) de la discipline. Réputés pour avoir reçu une formation aux médias (informatiques, audiovisuels et désormais digitaux), les diplômés en Sic sont perçus comme des spécialistes, dans les milieux de l'information et de la communication et dans les services de communication des organisations, particulièrement en ce qui concerne la conception et l'animation des outils de communication. Or ces milieux organisationnels sont de plus en plus marqués par l'automatisation du traitement des données et par les réseaux socionumériques, qui transforment les métiers tel que celui de documentaliste, mais aussi ceux de la communication médiatique (presse, radio, télévision), des agences de communication 360° et des services de communication dans les organisations.

Selon A. Casilli : « Les réseaux numériques, l'informatique ubiquitaire et la prolifération des bases de données affectent nos modalités de coexistence. Les usages technologiques actuels influencent non seulement les transformations des pratiques

1. LMD, http://www.enseignementsup-recherche.gouv.fr/cid20190/organisation-licence-master-doctorat-l.m.d.html.

communicationnelles et productives, mais également l'émergence de nouvelles définitions de soi, de nouvelles occasions d'interaction sociale, de phénomènes politiques inédits » (Casilli, 2011, p. 5). Dans les organisations, les pratiques professionnelles, hybrides ou « natives », s'ajoutent et se conjuguent aux pratiques digitales privées. Ce processus et ces transformations techno-humaines se déployant dans les environnements de travail forment ce que nous appelons les Humanités Numériques au Travail (HNT). Dès lors, la professionnalisation des communicateurs épouse la complexité de ce contexte en évolution constante.

Le chercheur Wittorski, dans la note de synthèse qu'il rédige sur le sujet en 2008 dans la revue *Savoir*, définit la professionnalisation comme le processus qui constitue des professions, les construit et les organise. Actuellement le numérique s'impose comme une composante du fonctionnement de toutes les organisations, avec l'usage des outils et dispositifs numériques d'information et de communication. La compréhension des dimensions, tant performatives que symboliques, de ces dispositifs est devenue très rapidement, et de façon incontournable, indispensable dans les formations en Sic. Actuellement, la maîtrise du numérique dans l'activité professionnelle va désormais au-delà de l'usage habile et intuitif des natifs. Elle entre dans un faisceau de compétences instituées et attendues. Wittorski, il y a plus d'une décennie, définissait la professionnalisation comme « une tentative d'articulation plus étroite entre l'acte de travail et l'acte de formation [...] et une tentative d'intégrer [...] dans un même mouvement l'action au travail, l'analyse de la pratique professionnelle et l'expérimentation de nouvelles façons de travailler » (2008, pp. 14-15).

Cette professionnalisation est désormais associée au numérique et caractérise les études et les métiers de l'information-communication. Jean-Claude Domenget et Alexandre Coutant décrivent bien cet état de fait dans la préface du numéro 3 (2015) de la revue *Communication & Professionnalisation* dédié aux compétences du « Communicateur bousculé par le numérique »[2]. Ils précisent que les formations des communicateurs évoluent avec l'informatique et que les sciences de gestion jouent un rôle majeur. Les formations se trouvent partagées entre l'apprentissage de la stratégie, liée à une compréhension plus globale des enjeux du numérique à tous les niveaux de l'organisation, et celui des techniques liées à la maîtrise des supports. Comme nous

2. https://resiproc.org/publications-du-resiproc/cahiers-du-resiproc/

allons le faire plus loin, les auteurs s'interrogent sur « la part à accorder aux compétences stratégiques, opérationnelles et techniques » (Domenget et Coutant, 2015, p. 9).

Le défi et les enjeux de la formation résident dans l'ajustement nécessaire entre les niveaux de compétence attendus et les qualifications proposées, entre les points de vue, les postures, les usages des technologies numériques par les diplômés et les attentes des responsables des organisations susceptibles de les employer. Cette question, qui mobilise nombre de chercheurs en Sic, a été abordée lors de la journée de la commission « Éducation et Formation » de la SFSIC (2017) consacrée à l'appropriation du phénomène des *datas* par les communicateurs et aux enjeux de formation suscités par ce phénomène. « Face à la concurrence des approches statistiques et informatiques, comment former nos étudiants aux *datas*, tout en restant dans le champ des Sic et des métiers de la communication ? » (Collet, Maas, 2018, p. 201), tel était le questionnement essentiel de cette journée. L'analyse d'une sélection d'enquêtes sur les compétences des professionnels de la communication en Europe nous permettra de revenir sur cette problématique qui concerne directement les Humanités Numériques.

Les recherches centrées sur les HN, telles que les Sic les envisagent actuellement, irriguent l'ensemble des cursus disciplinaires. En effet, penser la sphère des nouvelles technologies s'avère central dans les formations en Sic. Il s'agit d'intégrer dans la réflexion pédagogique la combinaison de la technique et sa pratique, les usages et les usagers de tous niveaux et de toutes proximités, les relations de coopération induites et les nouvelles relations hiérarchiques, la production de services et/ou produits innovants, autant d'éléments qui forment, produisent et régulent ces Humanités Numériques professionnelles.

Notre propos portera spécifiquement sur les phénomènes info-communicationnels qui façonnent les Humanités Numériques en environnement de travail. L'originalité de cette proposition est d'affirmer, comme nous l'avons suggéré plus haut, que se conjuguent des transformations numériques individuelles et collectives (vie privée et sociale) avec d'autres pratiques numériques cadrées par les milieux professionnels. C'est bien à leur complexité que renvoie l'appellation d'Humanités Numériques au Travail.

Montrer comment ce qui pourrait être un nouveau paradigme des HNT restructure la communication organisationnelle nous invitera, dans une première partie, à reformuler notre questionnement,

puis à présenter une partie des travaux des Sic en Communication Organisationnelle (CO) sur la transformation numérique des organisations. Pour amorcer notre propos sur le triangle organisation-numérique-professionnalisation dans les Sic et l'émergence des travaux sur les Humanités Numériques au sein des organisations, il nous paraît incontournable d'évoquer les activités du GER (Groupe d'Études et de Recherches) Org & Co de la SFSIC. Réunissant un grand nombre de chercheurs en communication organisationnelle, ce groupe a travaillé ces dix dernières années à la compréhension des relations à la technique et aux évolutions numériques dans des environnements professionnels variés. Nous rendrons compte des réponses que ces chercheurs ont su apporter aux défis des Humanités Numériques au travail (HNT) pour lesquelles une professionnalisation info-communicationnelle s'avère nécessaire, adaptative et ouverte.

Dans un deuxième temps, nous proposons de questionner la professionnalisation des communicateurs en mettant en perspective les HNT évoquées dans la première partie avec ce qui ressort de différentes études et baromètres réalisés par les associations professionnelles et de recherche (EUPRERA, EACD, Entreprise et Média…) sur le métier de communicateur. Nous identifierons les évolutions des métiers de la communication dans les organisations et en particulier les évolutions des compétences en lien avec les HNT telles que les envisagent les professionnels, tant communicateurs que managers. Pour ce faire, nous nous appuierons essentiellement sur trois études :

1. Le chapitre « Skills, Knowledge and Competency Development » de l'European Communication Monitors édité par l'European Public Relations Education and Research Association (EUPRERA) et l'European Association of Communication Directors (EACD)

2. le baromètre UDA-Entreprises et Médias « Communication d'entreprise » réalisé par Harris Interactive (2017).

3. L'étude Cision (2019) sur la place des Directeurs de la Communication dans l'entreprise.

Dans cette partie, nous analyserons la professionnalisation au regard de ces études, notamment les tensions entre une opérationnalité immédiate souhaitée par les organisations et une distanciation critique qu'offrent à la fois les recherches et les formations universitaires en Sic. Il se détache de fait deux axes épistémologiques que nous prenons en compte : l'un de nature fonctionnaliste, l'autre issu des approches critiques. On retrouvera ceux-ci dans les recherches

sur la professionnalisation des communicateurs, proposées entre autres par Valérie Lépine (2016). Nous mettrons en perspective les évolutions des compétences et connaissances des communicateurs avec les travaux du RESeau International de PROfessionnalisation des Communicateurs (RESIPROC)[3]. Nous mobiliserons le numéro 3 de sa revue *Communication et Professionnalisation* déjà évoqué et intitulé « Le communicateur bousculé par le numérique : quelles compétences à transmettre ? » (Coutant, Domenget, 2015).

Nous conclurons sur des Humanités Numériques profondément intégrées aux formations et aux recherches en info-communication organisationnelle. Nous formulerons enfin quelques préconisations d'ordre général qui répondent aux enjeux de la formation en Sic dans des organisations en mutation numérique.

La transformation numérique des organisations à l'épreuve de la recherche en Sic

La Société française des Sciences de l'Information et de la Communication (SFSIC) a accueilli en son sein des Groupes d'Études et de Recherches qui organisent une réflexion thématique sous des formes diverses et variées (journées, rencontres avec les professionnels, colloques, programmes de recherche…) dans les domaines qui constituent des sous-champs de la discipline des Sic. À ce propos, l'ouvrage très complet réalisé par les membres de la Conférence permanente des directeurs-trices des unités de recherche en sciences de l'information et de la communication (actuellement en ligne et en actualisation permanente)[4] offre un panorama quasi-exhaustif de la diversité des recherches de la discipline. Le GER qui traite du domaine de la communication organisationnelle s'est intéressé en particulier à la place du numérique dans la formation, en écho à celle dans les organisations, lors notamment de journées d'études dans les années 2010 et de deux colloques (Nice 2013 et Cerisy-la-Salle 2017). Ces démarches pro-actives de réflexion sur le numérique résultent de l'activité des enseignants-chercheurs en Sic responsables des formations, ainsi que des institutions d'enseignement et de

3. Le RESIPROC et sa revue Communication et professionnalisation qui cherche à « *comprendre les évolutions des pratiques en communication, d'interroger le rôle des formations universitaires en communication, de renforcer le dialogue entre les communautés professionnelles et universitaires et, in fine, de définir ce qu'il faut entendre par professionnalisation en et de la communication.* » https://resiproc.org/a-propos/

4. http://cpdirsic.fr/wp-content/uploads/2018/09/dynamiques-des-recherches-sic-web-180919.pdf.

recherche. En effet, celles-ci ont désormais intégré les enjeux liés à la formation aux Humanités Numériques, tels que les travaux du domaine des Sic centrés sur l'apprentissage et la formation peuvent l'attester. Les enseignants-chercheurs ont d'autres enjeux cruciaux qui infléchissent, organisent les formations destinées à professionnaliser les communicateurs, dont la numérisation de toutes les activités organisationnelles demeure au centre des préoccupations. C'est pour bien saisir la distance entre ces enjeux et les contenus des formations que le GER Org & Co de la SFSIC a organisé régulièrement, entre 2012 et 2018, des évènements qui ont invité les milieux professionnels à s'exprimer sur les effets sociaux de l'intégration du numérique dans les organisations et les transformations effectives qui s'ensuivent.

Alemanno et Gramaccia (2014)[5] proposent une synthèse de ces réflexions, exposées dans un ouvrage entièrement consacré à l'intégration du numérique dans les organisations et ses effets sur leur management. Les auteurs dégagent de l'ensemble de ces résultats une mutation des « faits organisationnels » qui entraîne de nouvelles formes de relations au travail liées à la numérisation généralisée des échanges et de la production. De façon générale, telle que l'ouvrage de 2014 permet de l'apprécier, une diffusion insistante du numérique, que ce soit dans les activités professionnelles ou dans l'usage des réseaux sociaux numériques et des objets connectés, est le produit de la représentation et de l'appréciation des dirigeants, en fonction de leur secteur de production. Les études consacrées aux Directeurs de la communication montrent en effet l'importance accordée par les dirigeants à ce qu'ils nomment « la transformation digitale de l'entreprise ». Cette situation installe des Humanités Numériques organisationnelles qui peuvent se définir comme l'utilisation d'outils d'information et de communication numériques pour la diffusion, le partage-échange des connaissances et la coopération dans l'activité professionnelle. Or, selon la précipitation, la massivité ou même la lenteur dans cette installation du numérique, les politiques managériales font prendre des risques à la structure et aux individus. Ces risques sont souvent gérés par les professionnels de la communication. Ces derniers s'attachent à nuancer et réguler certaines caractéristiques narratives managériales relatives à un hypothétique réenchantement par le numérique, et à prendre en compte la transition numérique dans laquelle se retrouve

5. Alemanno S., Gramaccia G. (2014) « Multi-éclairages conceptuels et pratiques communication organisationnelle-management et numérique des textes », in *Management et numérique*, Paris, L'Harmattan.

l'organisation. Ils sont le plus souvent confrontés à des inquiétudes, des résistances de la part des collaborateurs, voire du management en lien avec les contradictions entre les attentes et la réalité des HNT. Ces situations demandent aux communicateurs d'innover et de faire appel à des compétences auxquelles ils n'ont pas forcément été formés. À travers l'analyse de notre terrain, nous tenterons de faire émerger ces compétences essentielles à leur reconnaissance et à leur professionnalisation (Coutant, Domenget, *op. cit.*), puis de les mettre en perspective avec les propositions de formations telles que les envisagent les chercheurs en Sic.

Une professionnalisation qui ne se limite pas à une dimension technologique et irrigue l'ensemble de l'organisation

Dans cette partie, nous questionnons la manière dont les professionnels de la communication intègrent les Humanités Numériques au travail. En effet, il nous semble que les mutations numériques des organisations sont sous influence forte des stratégies de communication et des outils qui en découlent, lesquels sont principalement numériques. Nous nous intéresserons aux pratiques des professionnels, telles qu'ils les décrivent au travers d'études, donc selon leur dire et non selon des actes observés en situation.

Le numérique, et particulièrement Internet, auraient-ils à ce point envahi nos activités professionnelles que des métiers spécifiques en découlent ? C'est cette question, issue du baromètre Axa des préoccupations des chefs d'entreprise[6], qui introduit l'avant-propos du dossier déjà mentionné de la revue *Communication et Professionnalisation*. Nous analyserons ici, dans une perspective fonctionnaliste, les compétences des communicateurs spécifiques aux HNT, telles qu'ils les perçoivent et les décrivent. Pour ce faire, nous nous appuierons sur trois enquêtes quantitatives et une plus qualitative, qui nous semblent complémentaires au niveau des publics et d'un point de vue géographique international. Nous avons choisi ces enquêtes, car elles ne se limitent pas aux dimensions numériques des métiers des communicateurs, mais au contraire traitent de la professionnalisation d'une manière globale et en relation avec l'écosystème organisationnel qui participe de cette professionnalisation.

6. Baromètre AXA des préoccupations des chefs d'entreprises en 2015 http://www.ipsos.fr/comprendre-et-maitriser-son-marche/2015-09-22-2e-edition-barometre-axa-preoccupations-chefs-d-entreprise.

La première étude est européenne et traite des pratiques et compétences des communicateurs, il s'agit de l'*European Communication Monitors*. Conduite auprès d'environ 3 000 professionnels de la communication, elle est réalisée tous les ans depuis 2007 dans plus de 40 pays par *l'European Public Relations Education and Research Association* (EUPRERA) and *l'European Association of Communication Directors* (EACD). Nous nous intéresserons essentiellement au chapitre « *Skills, Knowledge and Competency Development* » de l'étude 2016 et aux évolutions des outils et des enjeux stratégiques de la communication dans les éditions de 2017 et 2019.

Nous compléterons notre aperçu par les résultats de deux études réalisées en France et consacrées plus spécifiquement aux directeurs de la communication. Il s'agit d'une part du baromètre UDA-Entreprise & Médias portant sur la « Communication d'entreprise ». Réalisée par Harris Interactive dans sa version 2017, cette étude quantitative analyse les réponses de 150 directeurs de la communication issus d'entreprises, d'administrations et de collectivités françaises. D'autre part, nous nous référerons à l'étude Cision[7], réalisée en 2019 auprès de 343 Directeurs de la Communication en France, essentiellement de grandes entreprises (+ de 500 salariés).

De l'analyse de ces études et à partir de notre propos sur le triangle organisation – numérique – professionnalisation, nous avons identifié trois niveaux de compétences relevant des HNT en lien avec l'évolution des missions des communicateurs et participant à la reconnaissance de la fonction. Il nous semble que ces niveaux, et les aptitudes qui leur sont inhérentes, sont apparus d'une manière chronologique pour les professionnels. Plusieurs schémas venant des études analysées viendront appuyer notre démonstration.

Dans une première phase, la maîtrise des outils de communication digitale paraît indispensable aux communicateurs pour être « opérationnels ». Sont-ils formés ? Comment ? Et jusqu'à quel niveau de compétences techniques ?

Dans une deuxième phase, la multiplication des métriques, liées à la consultation des réseaux sociaux et des données sur les organisations et leur public, permet aux communicateurs d'avoir accès à des outils de mesure de leur réputation et e-réputation mais également de pouvoir orienter leur campagne, leurs canaux et leurs messages en fonction des données sur leurs publics. Qu'en est-il des compétences

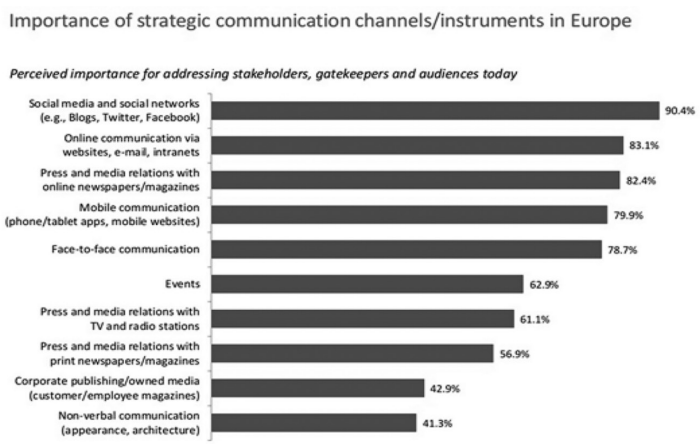
7. Cision est une entreprise de solutions et de services pour les RP, l'Influence, la Veille et l'Intelligence Médias destinés aux professionnels du marketing et de la communication.

nécessaires, à la fois techniques (maîtrise des outils de veille, des métriques liées aux consultations en ligne et aux *datas*), et stratégiques de compréhension du contexte, des enjeux, des différentes parties prenantes liées à l'organisation ? Jusqu'à quel niveau de technicité doit-on former les communicants qui ne sont pas des informaticiens ?

Dans une troisième et dernière phase, il convient de s'interroger sur l'accompagnement de l'évolution digitale des organisations et des modes de relations interpersonnelles liées aux HNT. Qu'en est-il des compétences et au-delà, des formations attendues ? Le rôle des enseignants-chercheurs en Sic n'est-il pas d'aller plus loin que les dimensions stratégiques et d'aborder les aspects déontologiques et éthiques de ces transformations (Collet, Maas, 2018) ?

La maîtrise technique des outils numériques « classiques » : un acquis ?

Nous nous intéresserons ici aux outils numériques tels qu'ils apparaissent dans les enquêtes. Le baromètre UDA montre ainsi que l'utilisation des outils dits « numériques » est devenue une norme pour les communicateurs. Les sites institutionnels, le fil twitter, les chaînes sur les réseaux sociaux, les pages LinkedIn et Facebook sont communément utilisés par plus de 7 entreprises sur 10. Selon l'enquête, 86 % des entreprises interrogées ont un fil twitter et 83 % utilisent une chaîne de vidéo en ligne pour leur communication *corporate*.



Graphique 1 - www.communicationmonitor.eu / Zerfass *et al.*, 2017

Q6: How important are the following methods in addressing stakeholders, gatekeepers and audiences today?

De même, l'*European Monitor* de 2017 liste les outils les plus importants dans le cadre du développement d'une communication stratégique selon les communicateurs européens. Le tableau illustre bien la prédominance des outils digitaux : pour 90 % d'entre eux, ce sont les réseaux sociaux qui permettent le mieux de communiquer avec les publics. Les outils autrefois traditionnels telles que les relations-presse avec les médias *off line* ne récoltent que 61 % des suffrages pour les télévisions et radios et 56 % pour la presse papier.

Ces études montrent en effet que les outils numériques les plus en vogue sont déjà devenus « classiques » par l'usage qu'en font les communicateurs. On constate également qu'il n'y a guère d'innovations dans le choix des canaux.

Les communicateurs sont-ils formés à une utilisation stratégique de ces outils ?

Ces outils numériques, pourtant déjà devenus une norme dans les directions de la communication aussi bien en France qu'en Europe, demandent des compétences techniques non nécessairement enseignées dans les formations universitaires. En 2017, d'après la cartographie réalisée par la Commission formation de la SFSIC, sur 187 mentions ou parcours de masters en Sic des universités françaises, seuls 36 font figurer « secteur du numérique » dans leur intitulé (Bessière, Collet, Corroy, Maas, 2017). La proportion est donc faible, même si cela ne signifie pas pour autant que les enseignements ne font pas référence aux outils numériques.

Or, selon les analyses, nous ne trouvons pas d'information sur le véritable degré de compétence technique dont font preuve les communicateurs dans la conception et l'utilisation de ces outils numériques, alors même qu'ils sont valorisés dans les discours.

Le tableau ci-dessous, qui procède également de *l'European Monitor*, renforce cette constatation. En effet, alors que l'offre des entreprises en formations sur la technologie demeure faible, les communicateurs souhaitent développer leurs compétences techniques.

Graphique 2 - www.communicationmonitor.eu/Zerfass *et al*., 2016

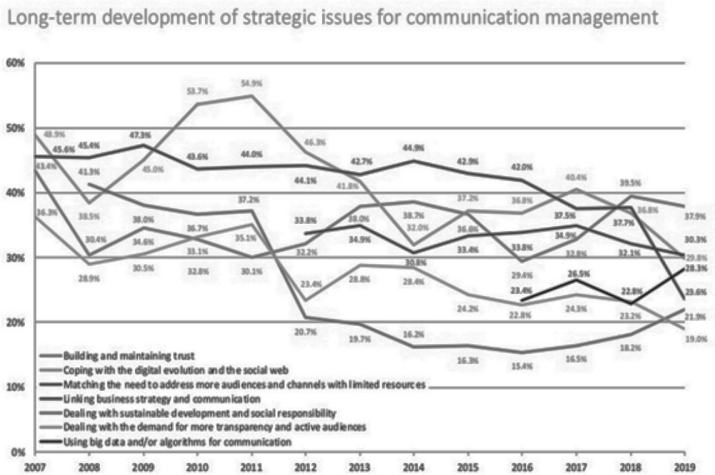
Q17: Thinking of yourself, your current capabilities and your future development, which of the following skills and knowledge areas do you believe are in need of developing?

Dans la même optique, l'étude Cision révèle que 60 % des directeurs de la communication interrogés aspirent, non à des augmentations de budget, mais « à des investissements dans de nouveaux outils/technologies (veille médias, plateforme d'influence, social media management, e-mail automation) ». 45 % d'entre eux souhaitent développer les compétences des collaborateurs dans cette direction. D'après cette étude, ils revendiquent ces investissements afin de « mieux digitaliser leurs process et gagner en efficacité ».

Cette difficulté de connaître le degré de compétence des communicateurs sur cette étape déjà ancienne de la transformation digitale des organisations est significative de fractures qui freinent la construction des Humanités Numériques au travail. À partir du moment où la maîtrise des outils numériques, et plus avant leur conception, participe de la professionnalisation des communicateurs, n'est-ce pas le rôle des formations académiques en Sic d'enseigner ces compétences techniques, bien qu'elles relèvent du graphisme, des sciences et techniques de l'information et de la communication, voire de l'informatique ? À quel niveau devrait se situer ce type de formation, licence ou master et sous quelle forme pour permettre aux étudiants d'acquérir un niveau de pratique suffisant ?

Datas et stratégie de communication : un mythe discursif face à la complexité de l'opérationnel

En revanche, les informations sur les *datas*, leur utilisation dans le cadre des organisations par les directions de la communication et les compétences associées à leur usage sont largement abordées par les études citées.



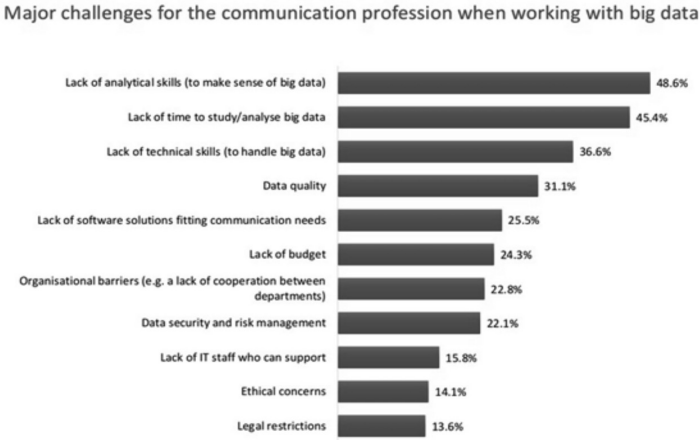
Graphique 3 - www.communicationmonitor.eu/Zerfass *et al.*, 2019

Q8: which issues will be most important for communication management?

Comme le montre ce tableau issu de *l'European Monitor* de 2019, toutes les stratégies de communication sont revues à la baisse par les communicateurs, sauf celles concernant la responsabilité sociale des entreprises et l'usage des *big datas* et des algorithmes. La volonté de manager par les *datas* la communication des organisations, que ce soit en termes de réputation, d'e-réputation ou d'influence, est clairement affichée. Cependant, ce tableau vient en opposition avec le baromètre UDA et la liste des outils réellement utilisés dans la pratique, lesquels demeurent plutôt « classiques ».

L'une des hypothèses développées par Alexander Kondratov dans ses travaux est que, au-delà des discours et du déclaratif, les *datas* sont dans la pratique peu utilisées dans les stratégies de communication (Kondratov, 2018). L'une des explications invoquées est « le manque de compétences nécessaires pour appréhender les données » (p. 125). Cette analyse est d'ailleurs confirmée par le tableau ci-dessous. Issu de *l'European Monitor* 2016 et illustrant les compétences des

communicateurs européens relatives aux médias sociaux, ce tableau tend à mettre en avant un niveau de compétences plutôt moyen quant à l'usage professionnel de ces outils banalisés.



Graphique 4 - www.communicationmonitor.eu/Zerfass *et al.*, 2016

Q 5: In your opinion, what are the three 3 major challenges for the communication profession in general when working with big data?

Ainsi, les communicateurs interviewés sont près de 50 % à déclarer un « manque de compétence pour analyser les *datas* et leur donner du sens », et 36 % « un manque de compétences techniques liées aux *datas* ». Alors même que ces compétences techniques, voire mathématiques, statistiques ou informatiques sont de plus en plus demandées aux communicateurs, quels rôles jouent les formations universitaires dans l'acquisition de ce type de compétences désormais « opérationnelles » (Coutant & Domenget, *op. cit.*) ? Quelle professionnalisation pour ces praticiens qui sont, d'après *l'European Monitor*, 56 % à déclarer avoir besoin de renforcer leurs compétences techniques ?

Il nous semble paradoxal de constater la prédominance des *datas* dans les discours des communicants des organisations mais aussi dans ceux des chercheurs et la rareté de leurs usages dans les pratiques organisationnelles, comme en atteste le graphique suivant.

Graphique 5 - www.communicationmonitor.eu/Zerfass et al., 2016

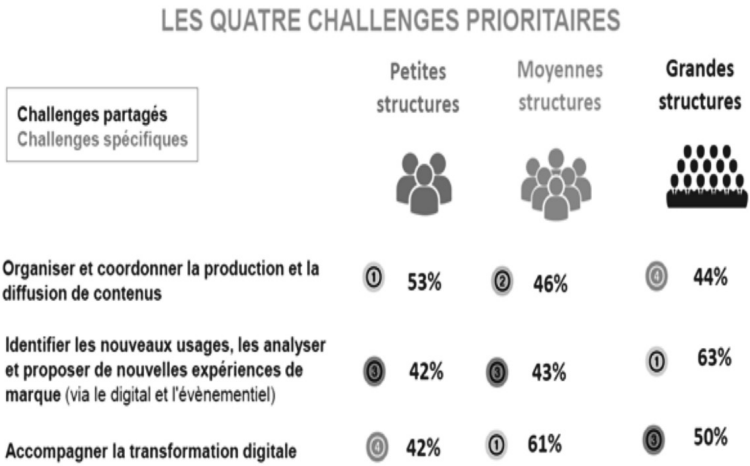
Q3: "Big data" is mostly described as huge volumes and streams of different forms of data from diverse sources (external and internal) and their constant processing, which provide new insights.

Alors que les outils numériques traditionnels et l'utilisation des réseaux sociaux sont véritablement plébiscités par les communicateurs dans leurs stratégies, nous constatons ici que seulement 21 % des organisations européennes avaient en 2016 implémenté des activités en lien avec les *datas*. Dans la version 2010 de ce moniteur, dans les objectifs de communication pour 2022, 28 % entendent développer l'usage des algorithmes et des *big datas* dans leur communication. Se pose alors la question d'intégrer aux formations en Sic des approches statistiques et informatiques afin de former les étudiants au *sourcing* et à l'analyse des données, légères ou massives. En effet, « ces données nécessitent un traitement pour être transformées en information » (Collet, Maas, p. 205, 2018) et, par là même, une certaine éthique de la part de ceux qui les analysent afin d'en assurer sécurité et confidentialité.

Dans le même temps, la notion de « transformation digitale des organisations », en lien avec les enjeux des directions de la communication, est largement convoquée dans toutes les études analysées.

Les enjeux actuels du communicateur : accompagner la transformation digitale des organisations vers de nouveaux modes de relations interpersonnelles

Que ce soit dans le baromètre UDA (graphique ci-dessous), dans l'enquête Cision ou dans *l'European Monitor* de 2019, accompagner la transformation digitale de leur organisation apparaît comme un défi majeur pour la majorité des communicateurs interrogés.



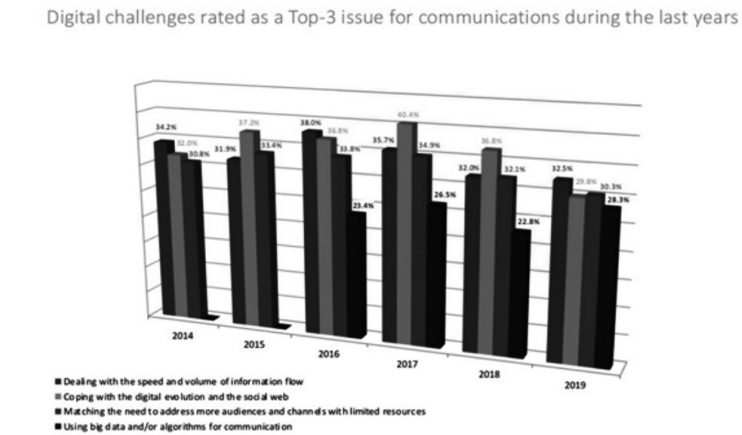
Graphique 6 - Présentation du baromètre UDA-Entreprises et Médias « COMMUNICATION D'ENTREPRISE » : P. 38

L'étude UDA indique également que les managers citent, parmi les risques pour l'avenir, des menaces liées au numérique, et en particulier l'échec éventuel de la transformation digitale de l'organisation et l'obsolescence rapide des compétences numériques.

Selon l'étude Cision, pour 54 % des Dircoms cette transformation n'est pas encore suffisamment à l'œuvre dans les entreprises, et 60 % d'entre eux sont en attente de meilleurs outils pour la parachever avec succès. Nous retrouvons ici notre questionnement concernant les compétences, mais aussi le manque d'innovation constaté dans les outils digitaux de communication communément utilisés dans les organisations. Nous pouvons également nous étonner de la vision purement fonctionnaliste, centrée sur les outils, de ces managers, lesquels gagneraient à intégrer une compréhension plus stratégique de cette dimension des HNT. De nouveau, on constate une contradiction entre les attentes et les réalités, en lien avec un

storytelling de réenchantement de l'organisation par le numérique. Le rôle des formations en Sic ne serait-il pas d'amener les étudiants, potentiellement futurs responsables de la communication des organisations, à aller au-delà et à appréhender les nouvelles formes de travail induites par la transformation digitale ? L'enseignement des approches critiques dans ces formations nous apparaît fondamental pour permettre aux communicateurs d'avoir une certaine réflexivité sur leurs pratiques.

L'*European Communication Monitor* de 2019 confirme cet enjeu pour les directions de la communication. Comme le montre le graphique ci-dessous, les évolutions digitales, les réseaux sociaux et l'utilisation des *datas* et des algorithmes dans la communication constituent depuis 2014 l'un des trois principaux objectifs stratégiques des communicateurs européens.



Graphique 7 - www.communicationmonitor.eu/Zerfaas *et al*. 2019

Q.8: Which issues will be most important for communication management, P. 56

Au-delà de l'analyse des chiffres, nous souhaiterions mettre en exergue deux points qui nous paraissent inhérents aux transformations digitales de l'organisation en lien avec les Humanités Numériques au travail. Le premier est que les inquiétudes des communicateurs concernant la transformation digitale sont largement infondées. Cette transformation a commencé dans les années 2000 avec l'avènement d'internet dans les organisations et ces outils numériques classiques et historiques dont nous avons vu qu'ils deviennent la norme. La transformation se poursuit et les organisations adaptent leurs outils. En revanche, les différences de temporalité dans le développement

des technologies à disposition et l'acquisition des compétences par les professionnels montrent que le véritable enjeu concerne les aspects sociaux, déontologiques et éthiques du développement de ces nouveaux modes de communications et leur appréhension par les collaborateurs.

En effet, la transformation digitale liée aux *datas* irrigue les pratiques professionnelles et les modes de collaboration. Ces aspects ne sont absolument pas évoqués dans les baromètres que nous avons analysés, alors que la compréhension des Humanités Numériques au travail englobe justement, au-delà des compétences techniques, des compétences de compréhension des enjeux sociaux et sociétaux, de réflexivité et une certaine éthique de la communication.

Ces études montrent que les techniques numériques ne cessent de se perfectionner, de s'affiner et de pénétrer de plus en plus l'activité humaine jusqu'à s'y substituer. La professionnalisation devient, dans cette mesure, de plus en plus complexe, car elle doit s'adapter au rythme des évolutions techniques et introduire dans sa stratégie cette accélération, se l'approprier et la faire adopter par les collaborateurs, en annoncer les transformations à venir et respecter les politiques managériales. Le rôle des formations en Sic dans la professionnalisation des communicateurs consiste alors à adapter les contenus en amont des évolutions des organisations, afin de leur permettre de mieux comprendre l'impact des technologies sur les Humanités Numériques au travail, notamment les médiations entre organisations et volonté managériale, entre attentes des individus dans leur vie professionnelle ou privée et attentes de la société.

Conclusion : questionner les Humanités Numériques par rapport à la professionnalisation, un enjeu majeur pour les Sic

La professionnalisation, qui s'associe à l'insertion professionnelle, est considérée comme l'une des missions principales de l'université. À tous les niveaux, stages imposés en DUT, BTS licence ou licence professionnelle, alternance possible pour nombre de masters en apprentissage ou contrat de professionnalisation, CIFRE ou autre contrat professionnel pour certaines thèses, sont autant de ponts utiles reliant le monde universitaire aux organisations publiques, privées ou associatives. Dans le champ des Sic, ce point de vue prévaut depuis l'origine de la discipline. En effet, les formations en Sic font partie des disciplines qui ne proposent pas de cursus académique de type

CAPES ou agrégation qui conduit exclusivement à l'enseignement. À leur issue, elles destinent les étudiants, pour une faible part à l'enseignement secondaire pour la documentation et l'audiovisuel, et à l'enseignement supérieur pour ces mêmes disciplines et toutes celles qui couvrent le champ des Sic. Mais le plus grand nombre d'étudiants trouveront leur voie en dehors de l'enseignement, dans les emplois de l'information (médias, journalisme…) et de la communication (industries culturelles, organisations…). Quelle que soit l'issue de leurs études, les communicateurs capitaliseront une relation à la recherche qui vient croiser les différentes formes d'expériences de terrain, humaines et technologiques. Le socle de compétences qui est à co-construire avec les professionnels de terrain vise à former aux *littératies numériques*[8], nouveau langage commun employé dans le cadre des communications professionnelles et personnelles.

Les recherches en Sic structurent et précisent donc aujourd'hui un portage pertinent des Humanités Numériques entre information et communication, instaurant ainsi un nouveau paradigme info-communicationnel au sein des SHS. Les différents domaines des Sic, dont l'info-communication organisationnelle, s'en sont emparés comme une évidence.

La professionnalisation des communicateurs basée sur les formations en Sic s'adresse essentiellement aux futurs professionnels qui entreront dans un univers organisationnel fait d'attentes managériales plus ou moins claires décrites plus haut. Pour conduire une transformation digitale dans une organisation, les Sic pourraient ainsi intervenir pour former l'ensemble des collaborateurs aux dimensions communicationnelles liées aux HNT. En effet, comme le rappellent Coutant et Domenget dans leur avant-propos : « ce changement implique à la fois un encadrement par la fonction communication mais aussi l'intégration par tous les employés de compétences de communication organisationnelle » (2015, p. 7).

Nos réflexions et études sur la professionnalisation des communicateurs par les formations en Sic impliquent enfin quelques remarques. D'une part, les communicants sont attendus par les directions générales sur la question du traitement des données, sachant que le métier de *data scientist* apparaît désormais dans les brochures des Ecoles d'informatique et IA de plus ou moins bonne qualité. Dans un avenir proche, les formations en Sic devront viser elles aussi la maîtrise du traitement des données par les organisations, afin de conserver une

8. https://primabord.eduscol.education.fr/qu-est-ce-que-la-litteratie-numerique.

visée critique sur leur utilisation. La manipulation technique des *datas* impose en effet une finesse éthique de leur traitement.

D'autre part, la transition digitale des organisations, impose de restructurer les formations, alors que la majorité des directions n'en comprend pas les enjeux ni en termes d'évolution du *core business* de l'entreprise, ni en termes d'évolution des formes de travail innovantes au sein des organisations, ni en termes déontologiques et éthiques. Nous rejoignons Manuel Zacklad lorsqu'il avance que « À la suite de Milad Doueihi (2008), nous considérons que les transformations introduites par le numérique induisent une évolution culturelle majeure, une conversion pour reprendre les termes de cet auteur, dont les implications anthropologiques sont multiples. » (Zacklad, 2015).

Enfin, l'apprentissage technique et technologique des outils numériques jusqu'au codage, voire la curation, devra être une base technique solide des connaissances du communicateur et le processus de professionnalisation en Sic gagnera à l'intégrer dans ses objectifs.

Les HNT conditionnent fortement la professionnalisation des communicateurs. Elles mettent les enseignants- chercheurs dans la nécessité d'adapter leurs enseignements pour repositionner plus fortement des connaissances numériques techniques et en même temps d'affiner leur posture critique dans leurs recherches dans l'objectif d'une percolation pertinente et utile de ces deux dimensions. Cet état de fait est dû aux avancées techniques assez peu prévisibles et l'obligation d'exploiter des *datas* dont le traitement intelligent, et surtout respectueux des individus, reste encore en jachère. Ce dernier point devient le défi majeur des HNT posé aux communicateurs.

POSITIONNEMENT

Les Humanités Numériques au Travail constituent le terreau de la professionnalisation des communicateurs, tout comme le travail de ces derniers est de contribuer à et d'accompagner la transformation numérique des organisations.

L'enseignement et la recherche universitaires assurent l'actualisation des connaissances numériques techniques et humaines, opérationnelles et critiques des communicateurs, nourries par les échanges avec le terrain des organisations de tout type.

Les formations aux compétences plus techniques telles que les statistiques ou l'informatique permettant l'analyse des *datas* doivent s'accompagner de formations à l'éthique des Humanités Numériques au travail afin de réguler les risques aussi bien pour les communicateurs que pour la société et de permettre à ceux-ci de développer des approches réflexives et critiques par rapport à certaines injonctions managériales et aux pratiques communicationnelles qui en découleraient.

BIBLIOGRAPHIE

Alemanno S., Gramaccia G., 2014. Multi-éclairages conceptuels et pratiques communication organisationnelle-management et numérique des textes, in *Management et numérique,* Paris, L'Harmattan.

Arvanitakis Sophie (2014) le bizprojet : une innovation technologique d'aide à la décision pour les patrons, in

Alemanno Sylvie P. (dir.), *Communication organisationnelle, management et numérique,* Paris, L'Harmattan, coll. « Communication et civilisation », pp. 91-100.

Bessière D., Collet L., Corroy L., Maas E. (coord.) (2018), « La cartographie des formations de niveau master de la SFSIC ». *Les Cahiers de la SFSIC n° 14*. Éd. SFSIC (Société Française des Sciences de l'Information et de la Communication), Paris.

Bourdoncle Raymond (2000) « Professionnalisation, formes et dispositifs ». In *Recherche & Formation*, N° 35, 2000. Formes et dispositifs de la professionnalisation. pp. 117-132.

Casilli Antonio (2011) « Présentation » in *Communications*, n° 88, « Cultures du numérique », pp. 5-8.

Collet L., Maas E. (coord.) (2018), « Journée formation de la SFSIC ». *Les Cahiers de la SFSIC n° 16*. Éd. SFSIC (Société Française des Sciences de l'Information et de la Communication), Paris.

Coutant Alexandre, Domenget Jean Claude (2015). Le communicateur bousculé par le numérique : quelles compétences à transmettre ?. Les Cahiers du RESIPROC, Presses universitaires de Louvain La Neuve.

De Ceglie Audrey, (2014), Les réseaux sociaux de femmes dans les entreprises : une nouvelle façon de percevoir le management organisationnel, in Alemanno Sylvie P. (dir.), *Communication organisationnelle, management et numérique,* Paris : L'Harmattan, coll. « Communication et civilisation », pp. 91-100.

Doueihi Milad (2008).*La grande conversation numérique*, Seuil, Paris.

Kondratov Alexander (2018), « Stratégies de résistances des professionnels des relations publiques à l'utilisation des *big data* dans les organisations en France et en Belgique » in « Pratiques de la communication et *big data* », Francine Charest et Anne Marie Cotton (dir.), *Communication et Organisation*, P.U.B.

Lépine, Valérie. (2016). Penser la professionnalisation comme une mise en mouvement : les communicateurs. *Revue française des sciences de l'information et de la communication*, 9 | 2016. http://rfsic.revues.org/2266; DOI : 10.4000/rfsic.2266.

Palakartcheva Marina (2014), L'impact de la visualisation de l'information sur le processus de prise de décision dans un contexte organisationnel, in Alemanno Sylvie P. (dir.), *Communication organisationnelle, management et numérique,* Paris, L'Harmattan, coll. « Communication et civilisation », pp. 123-137.

Sadin Éric (2013) *L'humanité augmentée. L'administration numérique du monde*, L'Échappée, coll. « Pour en finir avec », 2013, 160 p.

Sadin Éric (2015) *La Vie algorithmique. Critique de la raison numérique*, Paris, L'Échappée, coll. « Pour en finir avec », 2015, 288 p.

Zacklad, Manuel. (2014). « Humanités Numériques et digitalisation de la science », dans Actes du XIX[e] congrès de la SFSIC, Toulon, 4-6 juin, http://sfsic2014.sciencesconf.org/31853/.

Wittorski Richard (2008) « Professionnaliser la formation : enjeux, modalités, difficultés »,

WEBGRAPHIE

https://resiproc.org/publications-du-resiproc/cahiers-du-resiproc/

CHAPITRE XII

Cartographies et modélisation des formations en Humanités Numériques en France

Orélie Desfriches-Doria
Université Paris 8, Laboratoire Paragraphe
Elie Allouche
Direction du numérique pour l'Éducation,
Ministère de l'éducation nationale, de la jeunesse
et des sports, Ministère de l'enseignement supérieur,
de la recherche et de l'innovation (MENJ-MESRI)

Résumé : Dans cet article, nous donnons à voir le développement des *cursus* en Humanités Numériques au travers de la cartographie des contenus et discours associés à ces formations. L'objectif est d'effectuer un état des lieux des formations en Humanités Numériques. Après avoir présenté leurs caractéristiques et les sources disponibles sur leur répartition, une analyse des données ainsi extraites permet d'en dresser des cartographies et d'en proposer une esquisse de modélisation, avec un focus sur les articulations avec les enjeux de formation dans l'éducation.

Mots-clés : Humanités Numériques, cartographie, formations, métadonnées, modélisation.

Abstract : The aim of this study is to carry out an inventory of educational programs in digital humanities in France. The presentation of cartographies of contents and discourses associated to training programs' descriptions allows us to expose the development of programs in digital humanities. After presenting their characteristics and the available sources of data, an analysis of the data thus extracted makes it possible to draw up cartographies. Then we propose a model of these programs in digital humanities, and we connect on general issues in education.

Keywords : digital humanities, cartography, education program, metadata, modeling.

Tant dans la recherche que dans l'enseignement, les Humanités Numériques s'affirment à la fois comme un champ d'étude et comme une communauté de praticiens qui s'organise autour d'une approche interdisciplinaire des questionnements et méthodes scientifiques en lettres/arts/SHS et s'appuie sur les bouleversements épistémologiques apportés par l'utilisation des technologies numériques.

Une partie du travail présenté ici donne à voir la répartition non pas géographique des formations en Humanités Numériques (Humanités Numériques) en France, cartographie géographique déjà accessible à partir de la base de données Dariah. Il présente plutôt une vue globale des formations en Humanités Numériques à travers des visualisations des descriptions des formations et des métadonnées qui leur sont associées, à partir de données extraites de bases qui recensent des formations.

Dans une première partie nous présentons les caractéristiques globales des formations en Humanités Numériques La deuxième partie présente des cartographies des formations en Humanités Numériques en France par type de formation, par discipline, par objets et par techniques, d'après des données issues de la Dariah. Par la suite, une dernière cartographie des Licences d'origine des étudiants en Humanités Numériques, basée sur une extraction de données issues de « *Trouver mon Master* » est proposée. Pour chaque cartographie, les données sont présentées, puis les cartographies analysées. Enfin en troisième partie, nous proposons un modèle des grandes dimensions enseignées dans les formations en Humanités Numériques et ce, en s'appuyant sur l'observation d'un panel varié de formations à l'international.

Présentation et caractéristiques globales des formations en Humanités Numériques

Au cœur de ces formations se pose notamment la question de la connexion entre l'évolution des méthodes scientifiques, la

construction des savoirs, l'évolution des pratiques didactiques et l'évolution de la formation professionnelle. Nous constatons en effet une hybridation entre objets d'étude traditionnels, traditions de formation des institutions/unités de recherche et pratiques des technologies numériques.

Ces formations traduisent un effort d'articulation entre compétences techniques, compétences disciplinaires et démarche projet dans une optique le plus souvent interdisciplinaire. Cette articulation a été présentée par exemple en ces termes lors du THATCamp 2012[1] :

> « La tentative d'identification du digital humanist passe par une identification des compétences de base, qui seraient notamment la maîtrise des langages HTML, TEI et Javascript. » « Mais le répertoire de compétences du digital humanist dépasse le cadre de ces compétences techniques. Il est donc, de fait, difficile de trouver une base commune à l'ensemble des compétences des digital humanists. »[2]

Dans le même sens, nous pourrons constater, dans la partie suivante, la diversité des technologies mentionnées dans la description des formations en Humanités Numériques en France.

En effet, l'offre de formation en Humanités Numériques est d'abord fondée sur le rapport ancien de la science et de la technique et des conditions techniques de la construction des savoirs (Le Deuff, 2012).

Dans l'environnement numérique contemporain, les compétences mobilisées relèvent à la fois d'un effort de distanciation et de compréhension des mécanismes à l'œuvre nécessitant des compétences techniques (recueil et traitement des données, maîtrise d'outils, de normes), des compétences en culture numérique (littératie/translittératie), en design, informationnelles et organisationnelles (veille, documentation, travail collaboratif, gestion de projet, etc.).

---

1. Deux ans après la première session (durant laquelle avait été rédigé le Manifeste des digital humanities) le THATCamp Paris de 2012, organisé par le Centre pour l'édition électronique ouverte à l'École des Hautes Études en Sciences Sociales à Paris, fut une nouvelle occasion d'échanges et de partages d'expériences entre acteurs des Humanités Numériques, notamment sur les thèmes de la formation, des compétences, de l'organisation des communautés, de la diffusion et de l'instrumentation de la recherche.
2. Source : Berra A. et Le Deuff O., 2012. THATCamp Paris 2012 - Quelles compétences et littératies pour les Humanités Numériques ? – Éditions de la Maison des sciences de l'homme. In : THATCamp Paris 2012 : Non-actes de la non-conférence des Humanités Numériques [en ligne]. Éditions de la Maison des sciences de l'homme. Paris, Éditions de la Maison des sciences de l'homme. [Consulté le 29 mai 2019]. Disponible à l'adresse : https://books.openedition.org/editionsmsh/334.

L'analyse de ces formations confirme que les Humanités Numériques reposent à la fois sur des méthodes et compétences disciplinaires et transdisciplinaires (techniques, informationnelles, médiatiques), explorant des territoires déjà balisés ou de nouveaux territoires mis au jour par la mise en données numériques des savoirs, mais également sur des valeurs de partage et de diffusion des compétences et des connaissances (rappelées notamment par le Manifeste de 2010) (Dacos, 2010), et enfin, sur l'organisation d'une communauté de praticiens à dimension nationale et internationale.

Cartographie des formations en Humanités Numériques

Au cours de la présentation des visualisations de cette partie, il convient de garder à l'esprit que les données sont partielles et incomplètes, elles correspondent à un état à un instant T des données présentes dans les bases prises comme sources. D'autre part, certaines formations changent d'intitulé et de programme au fur et à mesure des évolutions des disciplines, mais également des technologies numériques, et ce, au fil du déploiement international des conceptions des Humanités Numériques Étant donné la quantité et qualité des données qui sont présentées et analysées dans cet article, nous tenons à préciser que les analyses qui ont été menées s'en tiennent volontairement à un niveau descriptif par souci de rigueur scientifique, afin de ne pas extrapoler des interprétations sur ces données imprécises, parcellaires, et qui posent de nombreuses questions procédurales, quant à leur production (pour la base de données Dariah notamment). Les données à propos des objets et des techniques sont des indicateurs complémentaires de celles sur les disciplines. Elles révèlent certaines visions et formes d'interprétations divergentes de l'objet des bases sources, de leurs objectifs, des formations décrites et de leur appartenance disciplinaire, qui est le lieu d'enjeux de prises de position, compte tenu de l'opacité du fonctionnement des financements nationaux de la recherche en France.

Les questions dans ce cadre concernent l'opportunité ou la possibilité de produire et de décrire les formations, à l'aide de référentiels de description des formations, et ce, en particulier dans le champ émergent des Humanités Numériques. Ainsi, nombreuses sont les définitions co-existantes des Humanités Numériques, elles-mêmes proposées par des enseignants chercheurs qui sont aussi les porteurs des formations étudiées ici, qui véhiculent donc des contenus et descriptions qui correspondent à leur vision des Humanités Numériques De plus, la description de ces formations dans les bases étudiées relève de

processus de communication à l'intention des futurs étudiants notamment. La question du niveau de complexité et de détail à adopter selon les publics et les stratégies de positionnement disciplinaire nous semble cruciale : dans les bases utilisées comme sources, les pratiques de description des formations sont hétérogènes. En effet, les formations repérées dans ces sources sont renseignées avec un nombre de mots clés associés qui varie fortement et avec des niveaux conceptuels eux aussi variables. Les stratégies d'indexation des formations (non explicites) nous semblent donc relever alternativement d'une volonté d'exhaustivité (avec un grand nombre de termes utilisés pour décrire une formation pour certains items) ou d'une stratégie d'indexation à visée prototypique (quelques mots-clés censés être représentatifs du contenu, pour en donner un aperçu).

Quel périmètre pour répertorier les formations en Humanités Numériques ?

Nous recensons ici les formations en Humanités Numériques qui s'affichent comme telles dans leur intitulé et qui sont répertoriées sur les portails présentés ci-dessous, ainsi que les sites des universités et unités de recherche. Pour établir une cartographie des formations en Humanités Numériques, au moins deux outils systématiques peuvent être utilisés. D'une part, la base de données produite par Dariah Course Registry[3] ; d'autre part, le portail national « Trouver mon Master ».

Autres outils pour s'informer

En premier lieu, la liste francophone de discussion autour des Digital Humanities (DH)[4] : « canal principal et historique des DH francophones (1 300 membres en 2018), cette liste de diffusion par mail est un espace dans lequel chacun(e) peut diffuser ses annonces, réflexions, etc. et recevoir celles des autres[5] ».

En second lieu, nous avons utilisé Calenda, le calendrier des lettres et SHS : « Calenda est un calendrier multilingue en sciences humaines et sociales en ligne et en libre accès. Lancé en 2000, il a vocation à permettre aux membres de l'enseignement supérieur et de la recherche d'être informés des événements scientifiques susceptibles de les intéresser. Il offre, grâce à ses URL pérennes, une archive

3. https://www.dariah.eu/tools-services/tools-and-services/tools/digital-humanities-course-registry/ (traduction modifiée à partir de Google traduction).
4. https://groupes.renater.fr/sympa/info/dh.
5. http://www.humanisti.ca/se-tenir-informe/.

de la vie scientifique dans le domaine des sciences humaines et sociales [...] Calenda est une réalisation d'OpenEdition développée par OpenEdition Center (USR 2004, anciennement le Centre pour l'édition électronique ouverte, Cléo-UMS 3287). Il est soutenu par Aix-Marseille Université, le CNRS, l'EHESS et l'université d'Avignon et Avignon Université. Il inscrit son action dans le cadre de la Bibliothèque scientifique numérique (BSN)[6] ».

Présentation des sources de données utilisées pour les visualisations

Nous présentons ci-après plus en détail les deux sources dont les données sont issues pour l'élaboration des cartographies présentées en partie 2.3.

— **Dariah Course Registry**[7] : il s'agit d'un recensement et d'une cartographie géographique[8] des formations proposées par l'infrastructure européenne de recherche numérique pour les arts et les sciences humaines (DARIAH). Dans cette base, la recherche peut s'opérer par les filtres suivants : par pays, par ville, par institution, par niveau. L'entrée dans les contenus de formation s'effectue quant à elle selon trois entrées : disciplines, techniques, objets.

Au 20 mai 2019, la base recense 209 formations en Humanités Numériques, dont 28 pour la France[9]. Pour la communauté francophone, le travail de recensement et de modération est assuré par le groupe « formations » de l'association *Humanistica*[10] (Association francophone des Humanités Numériques) qui « *propose de formuler des recommandations concernant les formations des étudiants et des jeunes chercheurs en Humanités Numériques* ».

— **Le service en ligne « Trouver mon Master »**[11], proposé par le site officiel du Ministère de l'Enseignement supérieur, de la Recherche

6. https://calenda.org/about.
7. Lien : https://registries.clarin-dariah.eu/courses/.
8. La présentation suivante a été faite par Aurélien Berra (chargé de la modération pour la France) dans un courriel du 05/10/18 sur la liste DH, accompagnée d'une invitation à compléter l'inventaire des formations : « Le DARIAH-CLARIN Digital Humanities Course Registry est un inventaire européen collaboratif des formations ayant une composante d'Humanités Numériques ». https://registries.clarin-dariah.eu/courses/. Si une formation à laquelle vous êtes associé(e) n'apparaît pas dans la base de données et sur la carte, il suffit de créer un compte et de saisir les renseignements pertinents. : https://registries.clarin-dariah.eu/courses/users/login.
9. https://registries.clarin-dariah.eu/courses/statistic.
10. http://www.humanisti.ca/groupes-de-travail/.
11. https://www.trouvermonmaster.gouv.fr/.

et de l'Innovation, présente l'offre de formation nationale au niveau Master. L'accès aux formations s'effectue par recherche libre ou selon une série de critères (mention, parcours/spécialité, établissement, localisation, périmètre géographique). Au 20 mai 2019 cette base recense 26 Masters en Humanités Numériques, en résultat à la simple requête « Humanités Numériques ».

Tableau 1. Comparatif quantitatif des données récoltées dans les deux sources

	Dariah Course Registry	Trouver mon Master
Nombre de résultats bruts en France	28	26
Nombre d'établissements différents	18	23
Nombre de Masters	13	23
Nombre de Parcours	/	57

Présentation et analyse des données du Tableau 1

La base de données Dariah recense des formations de Master, de Licence, et aussi des cours à l'international. Les chiffres sont issus d'une extraction de la BDD, basée sur l'unique filtre « France » appliqué au champ « Pays ».

Dans la base de données Dariah, l'entité de référence dans la structuration de la base est une formation, de quelque nature qu'elle soit dans la nomenclature. Certains établissements apparaissent plusieurs fois, dans la mesure où ils ont renseigné plusieurs formations dans cette base.

Le nombre de diplômes d'après cette extraction est de 13 Masters, auxquels il faut ajouter 2 Masters intitulés « recherche » qui n'entrent pas dans la même catégorie sur ce site, à cause de la dimension internationale de cette source. Les parcours différenciés et entrant sous le même intitulé de Master sont difficilement détectables et parfois non mentionnés. Ainsi, le Master « Humanités Numériques » de l'Université Paris 8 n'apparaît qu'une seule fois dans la base, bien qu'il compte 5 parcours différents en réalité. Les 13 autres formations se distribuent entre « Licence » et « Course », selon la nomenclature de Dariah.

« Trouver mon Master » ne recense que les formations de Master en France. L'entité de référence dans cette base est l'établissement d'enseignement, et dans chaque occurrence d'une fiche on trouve les différents parcours proposés sous le même intitulé de Master, avec des descriptions qui varient. Dans ce cadre, le passage de 26 résultats à

23 établissements s'explique par des doublons ou des fiches vides. Il y a donc en France **23** établissements d'enseignement **supérieur** qui proposent des Masters en Humanités Numériques, repérables dans cette BDD avec la requête « Humanités Numériques ». Chacune de ces formations compte un ou plusieurs parcours, et **57 parcours différents** sont actuellement proposés en France.

Il est important de préciser les éléments présentés ci-dessus, puisque ces données sont hétérogènes et structurées de manière différente, par conséquent, difficilement comparables. De plus, ces données restent incomplètes : il existe par exemple un Master dispensé à l'INA sur le patrimoine audiovisuel et numérique[12] qui n'apparaît dans aucune de ces bases de données, et qui n'est certainement pas un cas isolé.

Les outils de cartographie des formations

L'outil de visualisation utilisé pour mener ce travail est un instrument de cartographie dite « sémantique » (Desfriches Doria et Lavenir, 2017). Il s'agit de GraphCommons[13], outil en ligne permettant de faire des cartographies sous forme de graphes à petite échelle (1 000 nœuds max). Ses principaux avantages sont d'être gratuit, facile d'utilisation (interactions avec les graphes et les données *via* la carte ou des tableaux, facilité d'importation des données), et de pouvoir « embarquer » des modèles de connaissances, au sens de l'ingénierie des modèles, selon le principe des ontologies.

Statistiques et visualisations de la base de données Dariah

Pour la réalisation de ces visualisations, les données ont été extraites *via* l'API proposée sur la base de données Dariah et téléchargées, puis nettoyées et reformatées afin de les rendre importables dans GraphCommons.

Cartographie globale des formations par type de formation

Cette première cartographie[14], présentée en figure 1, établie d'après les données de la base de données Dariah, donne une vision globale

12. https://www.ina-expert.com/masters-audiovisuels-et-diplomes-de-deuxieme-cycle/diplome-ina-gestion-de-patrimoines-audiovisuels.html.
13. https://graphcommons.com/.
14. Consultable en ligne : https://graphcommons.com/graphs/9372006f-1640-47d5-b23d-1456d39d330f.

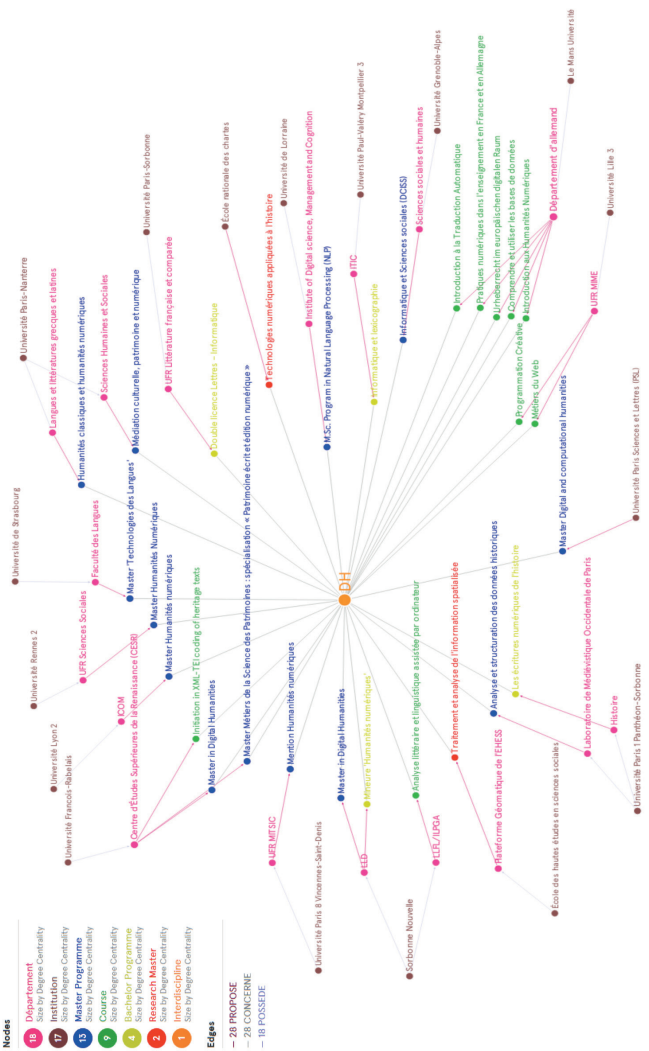


Figure 1. Cartographie globale des formations en HN en France, par type de formations

des formations recensées dans cette base (*Master/Master Recherche/ Bachelor Programme/Course*).

On peut observer d'emblée une surreprésentation des formations en Master (en bleu) dans cette cartographie globale. D'après les données de cette base, c'est donc le type de formation majoritairement proposé en Humanités Numériques en France. Ce résultat révèle une vision des Humanités Numériques qui correspondrait à une formation en double cursus : un cursus initial en Licence axé sur une des disciplines des humanités classiques, puis une formation en Master Humanités Numériques qui complète la première, selon le scénario le plus simple. On observe empiriquement, cependant, dans les publics étudiants des formations en Humanités Numériques, des parcours parfois plus complexes, avec des profils d'étudiants qui sont déjà titulaires d'un ou plusieurs Masters au moment où ils intègrent un Master Humanités Numériques. De véritables données chiffrées sur cet aspect sont difficiles à obtenir puisqu'elles relèvent d'un suivi des parcours d'étudiants à la charge des responsables pédagogiques. Ce suivi serait pourtant instructif à propos des profils des étudiants en Humanités Numériques qui constituent le vivier des futurs praticiens.

Cartographie par discipline

La cartographie[15] en figure 2 présente une visualisation issue d'une extraction de la base de données Dariah (à savoir les 28 formations recensées en France par cette base). Cette visualisation donne une vue globale axée sur la description des formations selon les disciplines renseignées pour chaque formation (**en rouge dans la figure** 2).

La cartographie est spatialisée avec l'algorithme « Force Atlas 2 », qui vise à rendre le graphe le plus lisible possible en optimisant l'espace afin qu'il y ait le moins possible de croisements entre les liens. Les nœuds les plus visibles sont mis en avant à travers un algorithme qui donne un poids relatif en fonction du nombre de liens entrants vers un nœud par rapport au plus petit nombre de liens entrants vers un nœud.

Analyse des données sur les disciplines

Il existe de fortes disparités dans ces données et elles suscitent des questionnements. Ainsi, les formations sont-elles décrites avec des mots-clés qui évoquent les disciplines dont les formations en Humanités

15. Consultable en ligne : https://graphcommons.com/graphs/525db979-df48-4d66-b370-49f7ca28f6a9.

Numériques se réclament, ou ces termes se réfèrent-ils à des disciplines connexes ? Par ailleurs, d'où vient cette nomenclature de disciplines, et par qui est-elle proposée ? Ces descriptions à base de mots-clés peuvent-elles être équivalentes aux données issues de « Trouver Mon Master » concernant les licences d'origine des étudiants préconisées pour entrer dans les formations en Humanités Numériques ?

Lorsque l'on télécharge ces données à l'état brut, chaque ligne correspond à une formation, et chaque discipline est classée dans une colonne ordonnée. Cela signifie-t-il que les disciplines sont ordonnées par ordre d'importance lors de leur saisie ? Cette interrogation questionne l'aspect contributif de la base de données Dariah, mais aussi l'encadrement des pratiques d'indexation et de description des formations.

En outre, les 28 formations sont décrites avec une fourchette en termes de nombre de disciplines mentionnées qui va d'un mot-clé à huit mots-clés pour décrire les disciplines. Or, dans le détail, sept formations ne présentent qu'un mot-clé, quand trois formations en proposent huit. Les formations qui ne sont décrites qu'à l'aide d'un seul mot-clé sont exclusivement de type « *Course* » ou « *Bachelor Programme* », ce qui pose la question de l'hétérogénéité des données en fonction du type de formation, et donc de la représentativité de la quantité des citations de mots-clés. Les Masters font l'objet de disparités dans la description, puisque leur nombre de mots-clés varie de deux à huit.

Enfin, au niveau de la généricité des mots-clés employés, que signifie l'utilisation d'un mot-clé « Social Science » alors que d'autres mots-clés plus précis existent dans les données comme « Histoire », « Archéologie », ou « Philosophie » ? Cette nomenclature est proposée pour décrire les disciplines au niveau international, mais peut-être existe-t-il des subtilités dans la considération de ces dernières, dans leur positionnement et même dans leur intitulé dans des contextes nationaux ?

Analyse de la cartographie par disciplines : les Sic en retrait

Dans la figure 2, les disciplines les plus citées pour décrire les formations du corpus ressortent visuellement.

On observe donc, d'après cette cartographie, la hiérarchisation en 5 niveaux dans la fréquence de citation des disciplines pour décrire les formations en Humanités Numériques en France. Ces rangs sont calculés en fonction du nombre de liens entrants vers les nœuds :

— Au premier rang (les nœuds les plus visibles et les plus grands) les disciplines les plus citées pour décrire les formations sont « *Literary and Philological Studies* » ;

— Au deuxième rang, les disciplines suivantes apparaissent : « *Linguistic and Language Studies* », « *Theory and Methodology of DH* », « *History* », « *Arts and Cultural Studies* », et « *Computer Sciences* » ;

— Au troisième rang, on observe les disciplines suivantes : « *Media and Communication Studies* », « *Human Language* » et « *Archeology* » ;

— Au quatrième rang apparaissent : « *Social sciences* », « *Cognitive Science* » et « *Library and Information Science* » ;

— Au cinquième rang, figurent enfin « *Musicology* » et « *Philosophy* ».

On peut donc constater que les Sic (« *Library and Information Science* » et « *Media and Communication Studies* ») apparaissent en retrait, puisqu'elles se situent en troisième et quatrième rang, dans ces descriptions conceptuelles des formations à l'aide d'intitulés disciplinaires. Cette sous-représentation des Sic peut aussi résonner comme un appel à contributions pour compléter les données disponibles dans la base Dariah.

Au-delà de ces observations, on remarque que les disciplines mentionnées sont nombreuses. Cette nébuleuse, qui se révèle étendue, renvoie aussi aux questions des fondations méthodologiques communes et des pratiques convergentes à travers cette diversité disciplinaire au sein des Humanités Numériques (Desfriches Doria *et al.*, 2018).

Cartographie par objet

La cartographie des objets des formations, présentée en figure 3[16], a été élaborée avec l'outil GraphCommons, et la spatialisation a été réalisée selon les mêmes paramètres que la carte précédente.

Analyse des données sur les objets des formations

La description des formations sur la base du vocabulaire des objets des formations est manifestement inégale quant à la quantité de mots-clés renseignés, engendrant donc une disparité sans doute trompeuse de ces objets dans la visualisation, à la manière de ce qui a été présenté précédemment.

16. Consultable en ligne : https://graphcommons.com/graphs/4bc0e2c1-58c0-470b-af78-1c45207abea1.

Figure 2. Cartographie des formations en HN en France, par disciplines

La nomenclature même de ces objets pose question, car certains objets désignent aussi des technologies. Ainsi, l'utilisation des termes « *data* », « *files* » ou « *tools* » ne peut pas être spécifique aux Humanités Numériques : nombreuses sont les disciplines ou les formations qui pourraient être décrites avec le même vocabulaire.

Par ailleurs, on retrouve ici des questionnements liés aux usages des langages documentaires et à celui de la visualisation proposée : la description des formations à l'aide des termes utilisés est-elle vue comme un système de descripteurs où chaque terme pris séparément fait sens précisément pour chaque item décrit, ou est-ce l'adjonction conjointement exhaustive des termes qui permet la description complète des items ? Ceci pose la question de la pertinence du traitement mené ici : le découpage en fonction de la fréquence des citations des mots-clés pour réaliser la cartographie. Le sens des termes utilisés, s'il se définit en référence à l'univers sémantique par proximité avec les autres termes utilisés pour un item, est perdu à l'occasion de ce découpage, par la suppression de l'environnement sémantique créé par les autres mots-clés utilisés pour décrire un item.

Analyse de la cartographie par objets : la « dominante implicite » du corpus

On observe, d'après cette cartographie, une hiérarchisation selon trois niveaux principaux dans la fréquence de citation des objets pour décrire les formations en Humanités Numériques en France :

— le premier niveau porte sur les objets « *Text* », « *DH* », « Data » ;

— le deuxième niveau mentionne « *Literature* » et « *Language* » ;

— le troisième niveau plus fréquent de citation indique « *Metadata* », « *Multimedia* », « Image », « *Manuscript* », « *File* » et « *Tools* ».

On peut remarquer la généricité des termes, ainsi que leur absence de précision. Une des dominantes observées porte sur le travail sur les données, métadonnées, les fichiers, le texte et plus globalement tout ce qui touche au traitement de corpus, sans que ce dernier terme soit inclus dans les objets cités.

Les Humanités Numériques apparaissent elles-mêmes comme objet d'étude, confirmant ainsi le côté réflexif des pratiques développées et la dynamique interdisciplinaire. Ce lien entre interdisciplinarité et dynamique réflexive a été étudié par certains articles récents (voir notamment Bénel, 2014). La notion de *projet*, fondamentale en Humanités Numériques (Desfriches Doria *et al.*, 2018), est cependant

Nodes

- 34 Objet — Size by metric_weighted_in_degree
- 17 Institution — Size by Degree Centrality
- 13 Master Programme — Size by Degree Centrality
- 9 Course — Size by Degree Centrality
- 4 Bachelor Programme — Size by Degree Centrality
- 2 Research Master — Size by Degree Centrality

Edges

- 194 PORTE SUR
- 28 DISPENSE
- 0 POSITIONNE

Figure 3. Cartographie des formations en HN en France, par objets

sous-représentée. Cet aspect renvoie à la question des pédagogies mobilisées dans les formations en Humanités Numériques. En effet, dans les Humanités Numériques, les projets en eux-mêmes déterminent en grande partie les compétences à mobiliser et à transmettre dans le cadre d'une formation, et non exclusivement l'origine disciplinaire des acteurs ou des institutions. Or, selon nous, les enseignements par projet en Humanités Numériques sont très profitables, notamment dans l'apprentissage du travail en collaboration interdisciplinaire, qui est une dimension forte de l'identité des Humanités Numériques (Desfriches Doria *et al*., 2018).

Cartographie des techniques

La cartographie des techniques[17], présentée en figure 4, a été élaborée avec l'outil GraphCommons, et la spatialisation a été réalisée selon les mêmes paramètres que les cartes précédentes.

Analyse des données sur les techniques

Le même type de disparités dans les données que précédemment présenté est observable, le nombre de termes pour décrire une formation du point de vue des techniques allant d'un seul terme à dix-huit.

Analyse de la cartographie par techniques : de la confirmation de l'importance des traitements appliqués aux corpus à l'ambiguïté des « frontières techniques »

On observe une structuration en plusieurs niveaux, comme dans les cartographies précédentes, pour la dimension « techniques » dans la description des formations, présentée dans la cartographie en figure 4. Les cinq niveaux proposés dans cette analyse sont regroupés par proximité des quantités de citations des termes, pour des raisons de praticité. Pour rappel, le nombre de citations d'un terme renvoie à la quantité de formations (entité centrale de la base de données Dariah) qui citent un des termes (par exemple « *Encoding* » ou « *Text Mining* »).

Le premier niveau est constitué d'un seul terme central pour décrire les techniques liées aux formations en Humanités Numériques, et il ressort fortement de cette cartographie comme le nœud le plus visible : il s'agit d'« *Encoding* » (cité 16 fois). Dans l'hypothèse où ces données sont fiables, la technique la plus représentative des

17. Consultable en ligne : https://graphcommons.com/graphs/51ef1758-d8d2-4dbe-917b-885f59583b82.

Humanités Numériques serait donc l'encodage. Cela semble cohérent avec la cartographie précédente qui révélait en filigrane l'importance du « *Corpus* ». Mais on peut s'interroger sur la nature générique de ce terme, qui est aussi utilisé en informatique, puisque toute donnée est toujours encodée, et sur l'absence de la mention de la TEI (*Text Encoding Initiative*) en tant que technique spécifique. Le terme est présent, mais dans un nœud de type « Intitulé de formation ».

Le deuxième niveau donne à voir les termes « *Searching* », « *Text Mining* » (cités 10 fois) et « *Information Retrieval* » ou encore « *Linked Open Data* » (cités 9 fois). L'emploi du terme « *Searching* » est ambigu : s'agit-il de recherche d'information, et quelle différence est alors proposée avec le terme suivant « *Information Retrieval* » ? En effet, dans la littérature spécialisée dans le champ de la documentation, ces distinctions prêtent à débat. Ces considérations qui relèvent de l'analyse des langages documentaires et de l'indexation, même si elles peuvent paraître complexes, empêchent toute généralisation sur la nature d'un éventuel consensus sur les techniques représentatives des Humanités Numériques.

Le troisième niveau présente les techniques suivantes : « *Concordancing* », « *Preservation Metadata* », « *POS-Tagging* » et « *Named Entity Recognition* » (cités 7 fois), et aussi « *Machine Learning* » et « *Topic Modeling* » (cités 6 fois). De nouveau, l'usage même de certains termes pose question : en effet, nous avons pu repérer le terme « *Text Mining* » au niveau précédent, mais l'utilisation du terme « *Concordancing* » qui fait partie des techniques de *Text Mining*, ainsi que la présence de « *Named Entity Recognition* » (« reconnaissance d'entités nommée », soit une composante des opérations réalisées par un programme de *Text Mining*) pose la question du niveau de précision des termes retenus pour décrire les techniques des formations en Humanités Numériques, et de l'utilité de cette démultiplication des termes en niveaux de détail enchâssés.

Le quatrième niveau concerne les techniques suivantes : « *Principal Component Analysis* », « *Cluster Analysis* », « *Commenting* » et « *Web Crawling* » (cités 5 fois). On peut noter que l'utilisation des techniques de *Crawling* est souvent liée à l'analyse par cluster, et cela renvoie au même type de questionnement que celui évoqué précédemment.

Le reste des nœuds se distribue comme suit :

— « *Open Archival Information Systems* », « *Collocation Analysis* » (4 citations) ;

— « *Mapping* », « *Pattern Recognition* » et « *Gamification* » (3 citations) ;

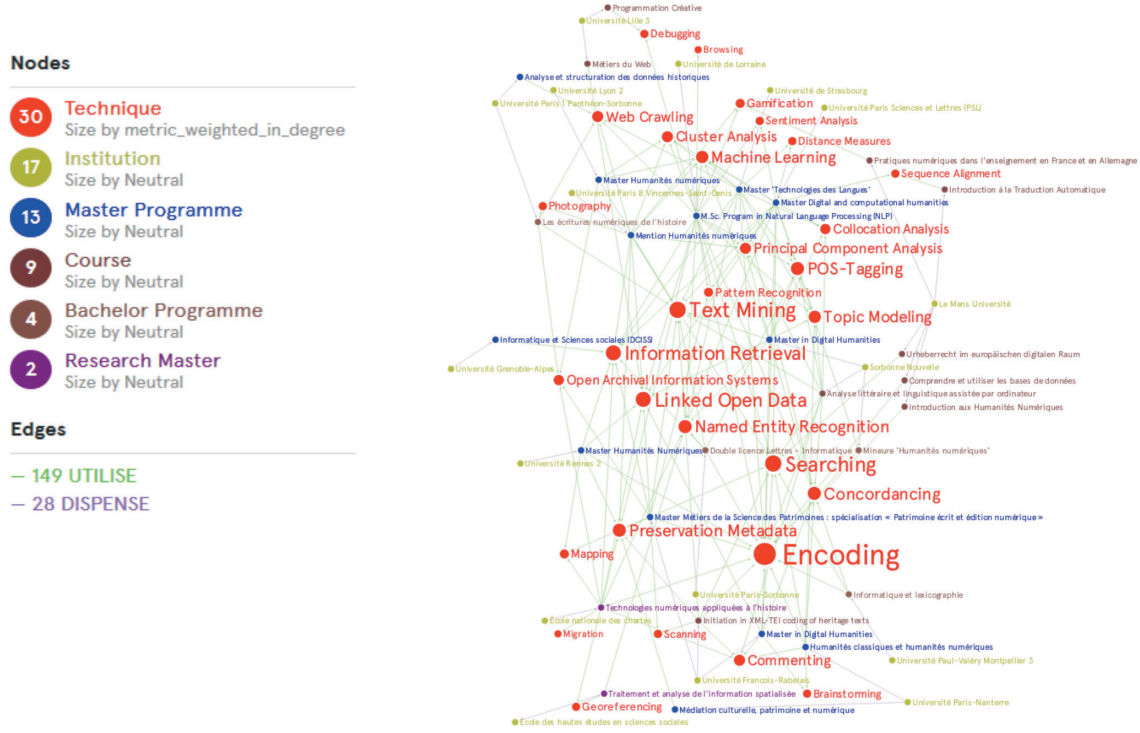


Figure 4. Cartographie des formations en HN en France, par techniques

— « *Scanning* », « *Sentiment analysis* », « *Georeferencing* », « *Photography* », « *Brainstorming* », « *Distance Measures* », « *Debugging* » et « *Sequence Alignement* » (2 citations) ;

— « *Migration* » et « *Browsing* » (1 citation).

Plus globalement, on constate une forte hétérogénéité des catégories de techniques incluses dans les termes pour décrire les formations. Ainsi, certains termes comme le « Linked Open Data » ou « Brainstorming » relèvent-ils bien d'une technique ? Il semble que le terme lui-même de « Technique » est suffisamment ambigu pour inclure tout autant une méthode de réflexion (*Brainstorming*), qu'une politique de publication, d'éditorialisation et de partage de descriptions et de liaisons de données (*Linked Open Data*).

La variabilité dans la spécificité des termes pose elle aussi question : est-il possible de mettre au même niveau des techniques de « *Machine Learning* » et d'« *Information Retrieval* » ? Que signifie l'emploi d'un terme comme « *Photography* » dans le champ « Techniques » ? Ce terme ne serait-il pas plus approprié pour décrire les « Objets » des formations ?

Cela pose les questions de la qualité des données utilisées pour réaliser ces visualisations et de la variabilité dans la précision des termes due aux aspects contributifs de l'indexation dans la base de données Dariah. Ces problématiques de cohérence inter-indexeurs sont une problématique classique en Sciences de l'Information-Documentation. De plus, une autre limite du travail présenté tient à l'identification des compétences visées à l'issue de ces formations, qui ne se réduisent pas nécessairement à des objets ou des techniques. Enfin, l'articulation du sens de ce champ « techniques » avec les formations n'est pas claire : est-ce que ce sont les techniques enseignées dans un programme formel d'enseignement, ou sont-elles pour certaines enseignées et pour d'autres évoquées, ou encore potentiellement mobilisées ?

La visualisation présentée en figure 5, qui bénéficie d'un autre paramétrage sur la base des mêmes données que la carte présentée en figure 4, permet d'affiner cette analyse, puisqu'elle présente les formations qui utilisent le plus de termes pour la description des techniques de leur formation. Les nœuds les plus visibles en termes de taille représentent les formations qui ont utilisé le plus de termes dans leur déclaration des données sur les techniques dans la base de données Dariah. On constate que les formations qui utilisent le plus de termes pour leur description sont majoritairement de type « Master » (en bleu dans cette cartographie qui suit).

Figure 5. Cartographie des formations en HN en France, par techniques, paramétrée en fonction du nombre de termes utilisés pour leur description

Approche chronologique de la création des formations en Humanités Numériques en France

En complément de l'analyse proposée ci-dessus, il est utile d'esquisser une approche chronologique de la création de ces formations en France. On observe une densification progressive de l'offre par type de formation. Les dates prises en compte correspondent à l'année de création de chaque formation recensée sur la base de données Dariah.

On peut constater que les deux premières formations créées en France datent d'avant 2010 et qu'elles sont de type « *Course* » et « *Research Master* ». Elles ont suivi la création de deux autres formations bien après 2010 de type « *Course* » et « *Master Programme* ». L'émergence de nombreux Masters en Humanités Numériques en France date de 2015 et d'après 2015.

Si l'on s'intéresse spécifiquement aux types de formations « *Research Master* » et « *Courses* », on constate une forte dimension instrumentale, technique, de ces formations concernant le type « *Course* » et un ancrage en Géographie et Histoire pour les « *Research Masters* ». La discipline de l'Histoire a été pionnière en matière de formation aux Humanités Numériques, puisque le premier Master recherche date de 2006 et s'inscrit dans cette discipline.

La majorité de la création des formations en Humanités Numériques concerne le niveau Master. En 2014, seul un Master existait. C'est donc très récemment que les programmes de Master en Humanités Numériques ont été créés en France. En 2016, il existait 6 Masters et 1 Bachelor. En 2018, on recense 4 Bachelors, et 13 Masters.

Traitement et analyse des données de « Trouver mon Master »

Les données qui servent de base à la réalisation de cette cartographie proviennent de « Trouver Mon Master ». Dans cette base, chaque établissement d'enseignement peut renseigner des détails sur chaque parcours de formation et notamment les licences d'origine des étudiants, considérés comme valides, ou préférées, ou encore préconisées, pour entrer dans ces parcours en Humanités Numériques

Cela apporte un éclairage sur le positionnement du parcours en termes de programme de formation et de positionnement dans le champ des Humanités Numériques

Cartographie des Licences d'origine

Les données ont été récoltées puis reformatées pour leur insertion dans l'outil qui a servi à la réalisation de cette carte. Le paramétrage a été réalisé selon les mêmes principes que ceux présentés précédemment.

Cartographie des licences d'origine : les Sic sur le devant de la scène

Avec cette cartographie[18] (Figure 9) on remarque que les Sciences de l'Information et de la Communication (Sic) sont en première position dans les formations dont sont issus les étudiants en Humanités Numériques en France. Viennent ensuite les Lettres, l'Histoire, la Philosophie et autres Humanités. En troisième position on retrouve l'Informatique, en quatrième les Langues (sous divers intitulés) puis en cinquième les Arts.

On peut opérer des rapprochements entre la cartographie des disciplines présentée plus haut et cette dernière. En effet, si la cartographie de la Figure 2 présente les disciplines dont se réclament les Humanités Numériques, ou ses disciplines connexes, celle des parcours d'origine des étudiants décrit les formations qui constituent le vivier des futurs praticiens des Humanités Numériques, qui feront évoluer les pratiques en Humanités Numériques au fil des promotions. On constate des disparités notoires. En effet, la place des Sic est en tête concernant les filières d'origine pour l'entrée dans les Masters Humanités Numériques, alors que dans la cartographie des disciplines les Sic apparaissent nettement en retrait. Il pourrait être judicieux d'interroger davantage ce résultat : est-ce l'appétence pour le numérique des étudiants en Sic qui en font un vivier favorable pour l'entrée dans des Master Humanités Numériques, ou bien la nature interdisciplinaire de cette section académique, ou encore l'intérêt pour l'écrit, le texte, la philologie, liée à la branche documentation des Sic ?

On observe également que les Arts se situent en 5e position dans la cartographie des filières, mais en deuxième dans la cartographie des disciplines. La Philosophie passe du dernier rang dans la cartographie des disciplines au deuxième dans la celle des filières d'origine des étudiants. Ces positions reflètent-elles un effet d'aubaine pour l'affichage de débouchés porteurs, pour les étudiants de filières dont on sait qu'elles sont difficiles en termes d'insertion professionnelle en France ?

18. Consultable en ligne : https://graphcommons.com/graphs/0c4c2eb5-5153-4f99-aff9-c7b9a6f08683.

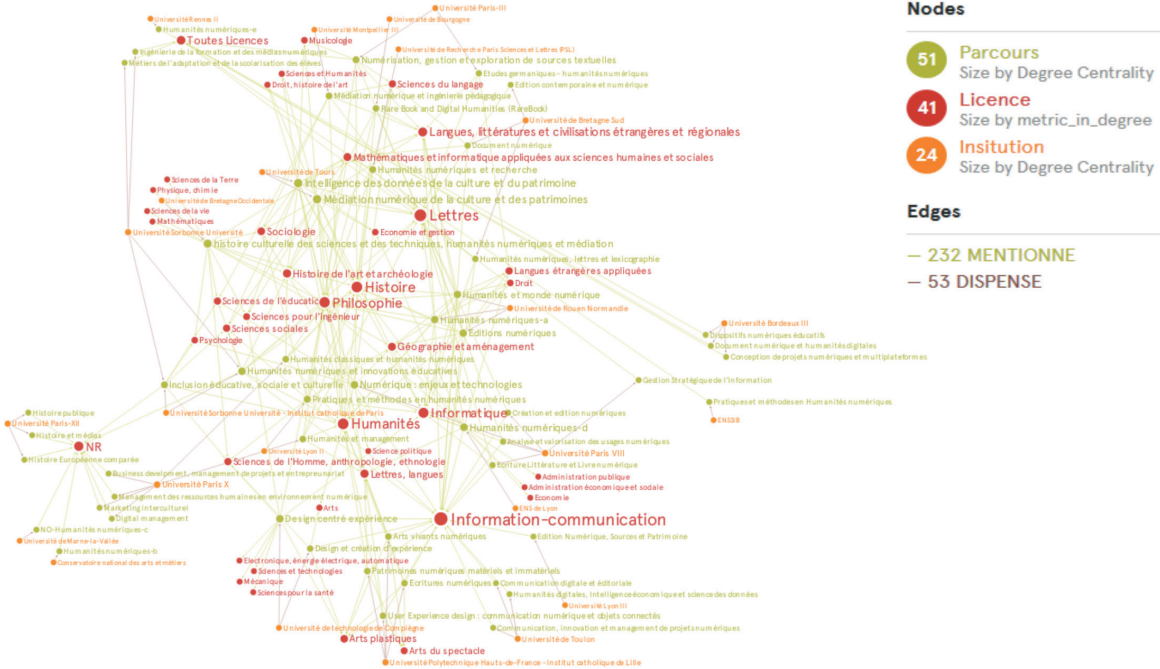


Figure 6. Cartographie des licences d'origine des étudiants admissibles en Master HN en France

Ces questionnements ne doivent cependant pas occulter les disparités dans les données rapprochées ici, et une requalification des disciplines mentionnées pour les filières d'origine des étudiants dans la nomenclature des disciplines de la base de données Dariah (ou l'inverse) permettrait d'affiner encore l'analyse.

Perspectives à propos de la description des formations en Humanités Numériques

Le travail présenté ici débouche sur de multiples interrogations : comment travailler à recenser des formations dans un domaine hétérogène tel que celui des Humanités Numériques, et comment proposer des langages documentaires de description de ces formations qui fassent consensus, dans un domaine émergent et en recomposition constante dans les contextes locaux de mise en œuvre et de projets ?

En outre, quelles sont les entrées pertinentes pour décrire ces formations ? Concernant les Humanités Numériques, est-il souhaitable d'opter pour une approche par compétences ? En effet, si pour les outils numériques et les technologies ces compétences sont repérables, elles apparaissent plus complexes à catégoriser dans les domaines liés aux humanités classiques.

Enfin, pour parfaire ce travail de cartographie du champ des Humanités Numériques à travers les formations, il nous semble qu'un prolongement de ce travail est nécessaire, sur le recensement des contenus de formations et leur décryptage à l'aide d'entretiens avec les responsables de ces formations, afin d'aller au-delà des discours institutionnels et de recenser la diversité des approches, des contenus de formations, et des visions des Humanités Numériques qui leur sont corrélés. Cette poursuite du travail sous forme d'enquête qualitative pourrait être complétée d'un travail maïeutique sur les trajectoires professionnelles des enseignants responsables de ces formations et sur leur profil académique.

Malgré la diversité des descriptions des formations présentées plus haut, nous proposons pour finir une esquisse de modèle des programmes de formations en Humanités Numériques.

Contenu et esquisse de modèle des formations en Humanités Numériques.

Pour analyser l'organisation et les contenus des formations en Humanités Numériques, nous faisons aussi le choix de répertorier des formations hors de France (essentiellement dans la communauté francophone et dans les pays anglo-saxons), des séminaires qui

proposent des rencontres entre chercheurs et autres acteurs de la communauté des Humanités Numériques et enfin des manuels ou autres supports accessibles en ligne.

En partant de l'analyse de plusieurs formations en Master ou doctorat (par exemple celle présentée par Claire Clivaz et Simon Gabay au séminaire « Épistémologie et pratiques des Humanités Numériques de l'Université Paris Diderot ») on peut aussi identifier plusieurs modules constituant une formation aux Humanités Numériques et esquisser ainsi une modélisation qui est présentée ci-après.

Généralités et apports théoriques dans l'enseignement supérieur :

- Aspects généraux et théoriques des Humanités Numériques
- Enjeux du numérique comme fait social
- Enjeux épistémologiques et didactiques du numérique et des méthodes computationnelles dans les pratiques disciplinaires et interdisciplinaires
- Environnement de travail et aspects organisationnels
- Environnement de travail, pratiques numériques et organisation de la recherche et du chercheur (incluant la question des réseaux sociaux, outils et plateformes de publication et de mise en réseau)
- Gestion de projet, généralement centrée sur une problématique de corpus
- Organisation du travail collectif (mise en réseau, mutualisation, coopération, collaboration) avec une ouverture internationale

Outillage technique :

- Compétences spécifiques requises pour le travail sur les données et métadonnées, leur utilisation pour la recherche et leur pérennisation (récolte/création, stockage, organisation, traitement manuel, traitement algorithmique, intelligence artificielle, visualisation…)
- Outils, normes et standards pour la création et le traitement des corpus
- Production, valorisation, diffusion
- Design, communication et valorisation
- Ouverture et diffusion des données et des savoirs : principes et pratiques du logiciel libre, de la publication en accès ouvert, des communs de la connaissance, etc.

Par ailleurs, plusieurs axes transversaux en matière d'**ingénierie pédagogique** peuvent être encouragés :

- L'utilisation de dispositifs de formation hybride, combinant l'utilisation d'outils de formation à distance et d'accompagnement

pédagogique en présentiel (cours, ateliers, moments d'échanges formels ou informels, etc.)

• Une approche didactique tournée vers une mise en pratique réflexive des concepts et outils

• Un effort de mise en cohérence entre modalités de formation et modalités d'évaluation, notamment par la production, la gestion de projet, la publication, la tenue d'un portfolio

• Une dynamique variable (hybridation, aller-retour ou une mise en tension) entre compétences disciplinaires et interdisciplinaires, d'où la question parfois soulevée de la pertinence d'une formation spécifique labellisée « Humanités Numériques » ou de compétences numériques transversales à inclure dans les blocs disciplinaires

• Un apport formatif spécifique issu du dialogue interdisciplinaire et interprofessionnel[19] inhérent aux projets en Humanités Numériques (enseignants-chercheurs, ingénieurs de recherche, bibliothécaires, documentalistes, designers, informaticiens, etc.) et de la place centrale des lieux de savoir (universités, centres de recherche, laboratoires, musées, bibliothèques, etc.) et des tiers-lieux.

Humanités Numériques et enjeux de formation dans l'enseignement scolaire[20]

Nous mentionnons ici une réflexion conduite depuis six ans sur les apports possibles du champ des Humanités Numériques pour la formation des enseignants, l'enseignement scolaire et l'évolution des pratiques pédagogiques. Cette réflexion est notamment développée dans le cadre d'un carnet de recherche dédié (Elie Allouche, 2014).

Le point d'étape qui peut être effectué ici peut rendre compte d'une part de l'intérêt de nombreux acteurs de l'éducation pour ce champ et la façon dont il peut alimenter leur réflexion et leurs pratiques, d'autre part d'initiatives multiples (rencontres, séminaires, formations) pour que cet intérêt se traduise concrètement par un réel travail collectif des enseignants sur l'évolution de leurs pratiques professionnelles dans une approche transdisciplinaire et ouverte sur la recherche.

Plusieurs jalons peuvent déjà être mis à l'actif de ce dialogue entre Humanités Numériques et éducation, parmi lesquels :

19. Dans la pratique cette interdisciplinarité se traduit soit par le couplage entre une discipline et le champ des Humanités Numériques, soit par le couplage de deux Masters (par exemple à Rennes, où le Master en Humanités Numériques vient en appui méthodologique)
20. Source de l'encadré : Elie Allouche, carnet de recherche « Numérique et éducation ». https://education.hypotheses.org/.

- les séminaires Tech day du Pôle numérique de l'académie Créteil sur « les Humanités Numériques à l'École » (22/06/16 et 21/06/17);
- le colloque international EdCamp « les Humanités Numériques pour l'éducation » (1er et 02/09/16);
- le « dialogue des Humanités Numériques » sur le salon Eduspot (08/03/17);
- le séminaire sur les Humanités Numériques dans le cadre du plan national de formation (26/03/18) ;
- création d'un groupe thématique numérique sur les Humanités Numériques par Direction du numérique pour l'Éducation (juillet 2020).

Conclusion

Dans ce chapitre nous avons présenté une analyse des formations en Humanités Numériques en France, selon plusieurs angles. Un premier angle dresse les grandes lignes des caractéristiques de ces formations. Le deuxième angle propose des visualisations de données à propos de ces formations. Ce travail de spatialisation relève lui-même des Humanités Numériques, confirmant une fois de plus la dimension réflexive de ce champ de recherches, hétérogène, mais riche de sa diversité. Il est important de souligner que les données utilisées correspondent à un état des descriptions des formations au moment de la collecte des données utilisées. Ces données sont, par ailleurs, issues de recherches dans des bases, qui ont été formulées avec des termes simples, elles sont donc potentiellement non-exhaustives. Ces données peuvent également évoluer dans le temps, puisqu'elles sont renseignées de manière contributive et déclarative. Il reste intéressant de constater les difficultés liées à l'hétérogénéité des descriptions et des pratiques d'indexation des objets étudiés, les formations en Humanités Numériques, qui, par définition, sont vouées à évoluer.

Un travail complémentaire d'enquête qualitative à travers des entretiens avec les enseignants-chercheurs qui élaborent ces formations et leurs programmes nous semblerait d'un grand intérêt, ainsi qu'une comparaison des données à l'international, notamment pour les aspects chronologiques. Concernant la description des formations sur la base de données Dariah, un travail comparatif des données à l'international permettrait d'observer l'existence hypothétique de caractéristiques nationales dans l'appréhension de ce champ émergent des Humanités Numériques.

Enfin, la troisième partie propose un modèle des contenus de formations, qui relève d'un travail de synthèse portant sur une diversité de sources, et incluant la question de la convergence avec des enjeux de formation pour l'enseignement scolaire. Cette proposition n'a pas de caractère prescriptif, mais pourrait constituer un support de dialogue pour les échanges entre les chercheurs, notamment dans la discipline des Sic, qui élabore actuellement son positionnement vis-à-vis des Humanités Numériques.

POSITIONNEMENT

La discipline des Sic n'est pas toujours visible à l'international, pourtant l'étude des données conduite ici révèle une diversité qui témoigne de la richesse des circulations entre Humanités Numériques et Sic au niveau national. Ce constat de la diversité des appréhensions des Humanités Numériques traduite dans les descriptions des formations est à l'image de l'originalité, de la singularité et de la (relative) liberté avec laquelle la recherche est menée en France, recherches auxquelles les formations en Humanités Numériques sont adossées. Ainsi, loin d'une vision consensuelle des Humanités Numériques, la diversité des approches et des profils des enseignants-chercheurs en Humanités Numériques en France fait écho à la diversité existante au sein de la discipline des Sic. Du point de vue international, les Sic et les SHS françaises bénéficient de particularités méthodologiques, et d'une forte tradition théorique. Aussi, les aspects stratégiques du positionnement international des Sic, et du positionnement national et international vis-à-vis de la communauté des Humanités Numériques appellent-ils selon nous à poursuivre ces dialogues intra-disciplinaires et interdisciplinaires. Cette circulation et ce dialogue à mener de manière continue et itérative, constituent selon nous un moteur à même de renouveler le dialogue intra-disciplinaire entre les branches communication et information des Sic, notamment en raison des dimensions techniques, méthodologiques et épistémologiques que ces débats partagent.

BIBLIOGRAPHIE

Allouche E. 2014, « Numérique et éducation : carnet dédié aux enjeux du numérique dans l'éducation ». Numérique et éducation. [Consulté 10 septembre 2019]. Disponible à l'adresse : https://www.openedition.org/16649.

Anon. 2000. *Calenda* – Le calendrier des lettres et sciences humaines et sociales dans Calenda – Le calendrier des lettres et sciences humaines et sociales [en ligne]. 2019 2000. [Consulté le 29 mai 2019]. Disponible à l'adresse : https://calenda.org/

Anon. 2014. *Digital Humanities Course Registry* dans Dariah Clarin Digital Humanities Course Registry [en ligne]. 2019 2014. [Consulté le 29 mai 2019]. Disponible à l'adresse : https://registries.clarin-dariah.eu/courses/.

Anon. 2014. Humanistica : Association francophone des Humanités Numériques dans Humanistica | Association francophone des Humanités Numériques [en ligne]. 2019 2014. [Consulté le 29 mai 2019]. Disponible à l'adresse : http://www.humanisti.ca/.

Anon. 2017. Le portail national des Masters – Trouver mon Master. In : Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation – Trouver mon Master [en ligne]. 2019 [2017]. [Consulté le 29 mai 2019]. Disponible à l'adresse : https://www.trouvermonmaster.gouv.fr/.

Benel A. (2014). Quelle interdisciplinarité pour les « Humanités Numériques » ? *Les Cahiers du numérique*, 10(4), pp. 103-132.

Berra A. et Le Deuff O., 2012. THATCamp Paris 2012 – Quelles compétences et littératies pour les Humanités Numériques ? Éditions de la Maison des sciences de l'homme. In : THATCamp Paris 2012 : Non-actes de la non-conférence des Humanités Numériques [en ligne]. Éditions de la Maison des sciences de l'homme. Paris, Éditions de la Maison des sciences de l'homme. [Consulté le 29 mai 2019]. Disponible à l'adresse : https://books.openediton.org/editionsmsh/334.

Dacos M., 2010. « Manifeste des Digital humanities ». THATCamp Paris. [Consulté 26 avril 2016]. Disponible à l'adresse : http://tcp.hypotheses.org/318

Desfriches-Doria O., Sergent H., Tran F., Haettich Y., & Borel J. (October 2018). What is Digital Humanities' identity in interdisciplinary practices ? : An experiment with digital tools for visualizing the francophone DH network. In *Proceedings of the 2nd International Conference on Web Studies* (pp. 39-47). ACM.

Desfriches Doria O., Lavenir C. (2017), « Document et Annotation : le cas des cartographies d'informations », dans M. Hassoun, K. Zreik, O. Larouk, G. Besacier (dir.), Actes du 20e Colloque International sur le Document Électronique – CiDE.20 : « Le document ? », 23-25 novembre 2017, ENSSIB, Lyon, France, pp. 297-308.

Drucker J., 2013. Introduction to Digital Humanities | Concepts, Methods, and Tutorials for Students and Instructors. In : [en ligne]. 2013. [Consulté le 23 avril 2016]. Disponible à l'adresse : http://dh101.humanities.ucla.edu/.

Kaplan F., 2013. DH101 2013/2014 course1 – Presentation of the course/Collaborative… In : [en ligne]. Lausanne. 2013. [Consulté le 29 mai 2019]. Disponible à l'adresse : https://fr.slideshare.net/frederickaplan/dh101-20132014-course1.

Le Deuff O., 2012, Humanisme numérique et littératies. *Semen - Revue de sémio-linguistique des textes et discours*, 34, 117. [Consulté le 16 avril 2020]. Disponible à l'adresse : https://archivesic.ccsd.cnrs.fr/sic_00776805

CHAPITRE XIII

Les enjeux de la formation de formateurs aux Humanités Numériques

Laurent Collet
LERASS, Université Paul Valéry Montpellier 3, France
Thibaud Hulin
ELLIADD, Université de Franche-Comté, France
Émilie Remond
Université de Poitiers, Techné, F-86073, Poitiers, France

Résumé : Face au développement de nouvelles offres de type « Humanités Numériques », le rôle du formateur est à interroger sous de multiples aspects. Comment peut-il se former ? Quelle vision les formations existantes proposent-elles ? Quelle place est réservée aux sciences de l'information et de la communication dans ces formations de formateurs ? Notre étude est exploratoire et vise à dresser un premier panorama des formations de formateurs aux Humanités Numériques. Nous proposons donc ici un premier inventaire des programmes de formation, en nous appuyant sur l'exemple d'offres existantes : masters « Humanités Numériques », formations au numérique dans les MEEF et formation des Unités Régionales de Formation à l'Information Scientifique et Technique (URFIST). Nous avons pu constater qu'il n'existe pas de formation de formateurs aux Humanités Numériques à proprement parler. Mais des modules se retrouvent dans les masters « Humanités Numériques », ceux des INSPE et dans les formations de l'URFIST. Les premières visent à former les professionnels accompagnant la transformation numérique des organisations, les secondes à intégrer les technologies numériques dans la pédagogie, les troisièmes à traiter plus efficacement les données. Face aux enjeux multiples de la numérisation de la société, les sciences de l'information et de la communication, discipline frontière, bénéficient d'un espace où développer une offre de formation de formateurs articulant habiletés numériques, esprit critique et démarches créatives.

Mots-clefs : Humanités Numériques, formation de formateurs, outils numériques, approches critiques et créatives, URFIST, INSPE, Master.

Abstract : Faced with the development of new training offers of the "digital humanities" type, the role of the trainer is to be questioned in many aspects. How can it be formed? What vision does existing training offer? What place is reserved for the information and communication sciences in these training courses for trainers? Our study is exploratory and aims to draw up a first panorama of training for trainers in digital humanities. We therefore offer here a first inventory of training programs, based on the example of existing offers: "digital humanities" masters, digital training in MEEFs and training of Regional Training Units for Scientific and Technical Information (URFIST). We have seen that there is no training of trainers in "digital humanities" per se. But modules are found in the "digital humanities" masters, those of INSPE and in URFIST training. The first one aims to train professionals accompanying the digital transformation of organizations, the second to integrate digital technologies in education, the third to process data more efficiently. Faced with the multiple challenges of the digitization of society, the information and communication sciences, a "liminal" discipline, therefore have a space where to develop a training offer for trainers combining digital skills, critical thinking and creative approaches.

Keywords: digital humanities, training of trainers, digital tools, critical and creative approaches, URFIST, INSPE, Master.

À un nouveau support d'écriture succède une nouvelle pédagogie. Peu après l'invention de l'imprimerie, les manuels se sont multipliés (Serres, 2011) comme étant un préambule nécessaire à l'émergence du Siècle des Lumières. Suite à l'invention de l'écriture numérique, déferlent aujourd'hui des MOOC et d'autres plateformes permettant la diffusion des connaissances à grande échelle, véhiculant des promesses d'un accès renouvelé et démocratique au savoir dans la perspective d'un nouvel âge éducatif. Les liens entre connaissances et technologies numériques tissent de nouveau la figure d'un savant polyvalent, à la fois lettré et inventeur. Cependant, cette vision idéale se heurte à des inquiétudes légitimes quant aux évolutions des formations : vision utilitariste du savoir, devenu dépendant de l'économie de marché et de la volonté des grands groupes industriels[1], logique concurrentielle dans les offres aux dépens d'un curriculum éprouvé : de telles craintes sont susceptibles de développer des incompréhensions entre les acteurs des sphères publiques et privées, mais aussi entre les disciplines scientifiques. Le thème de la *société de l'information apprenante* apparaît alors moins comme un objectif, que comme un levier idéologique qui s'inscrit dans le cadre d'une industrialisation continue de la formation (Moeglin, 2016).

1. Dans le secteur numérique, les groupes industriels leaders les plus connus sont les GAFAM (Google, Apple, Facebook, Amazon et Microsoft), auxquels il convient d'ajouter les NATU (Netflix, Airbnb, Tesla et Uber), les BATX (Baidu, Alibaba, Tencent, Xiaomi), etc.

Dans ce contexte, le formateur est l'interlocuteur privilégié d'apprenants toujours plus alphabétisés au regard des technologies numériques, sans pour autant qu'il s'agisse de « lettrés du numérique » (Bouchardon, 2014). Médiateur des apprentissages, le formateur accompagne les apprenants par la parole, mais aussi en s'aidant de ressources qui visent à renforcer son efficacité pédagogique (Moeglin, 2017). Face au développement de nouvelles offres de formation de type « Humanités Numériques », le rôle du formateur est à interroger sous de multiples aspects. Comment peut-il se former ? Quelle vision les formations existantes proposent-elles ? Comment les formateurs peuvent-ils assumer leur rôle ? Quelle place est réservée aux sciences de l'information et de la communication dans ces formations de formateurs ?

Notre étude est exploratoire et vise à dresser en un panorama des formations de formateurs aux Humanités Numériques en 2020-2021. Nous proposons donc ici un premier inventaire des programmes de formation, en nous appuyant sur l'exemple d'offres existantes : masters « Humanités Numériques », formations au numérique dans les MEEF et formation des Unités Régionales de Formation à l'Information Scientifique et Technique (URFIST). Nous poursuivrons par une réflexion sur les rôles à donner aux formateurs en contexte de formation, ainsi que sur les compétences et savoir-faire associés.

Les Humanités Numériques soutiennent l'éducation et la formation

Les technologies numériques peuvent être définies comme la forme que prennent certains objets-médias et qui peuvent être décrites à l'aide des mathématiques car elles sont par essence programmables ou calculables (Manovich, 2002), discrètes et manipulables. À l'inverse, la culture numérique renvoie à un ensemble de participations, de remédiations et de bricolages (Deuze, 2006) en situation d'interactions avec des objets numériques, constitutifs d'une identité et d'une histoire humaine. La société du numérique désigne alors un ensemble de paradigmes et de représentations définissant une société moderne en prise avec des savoirs, des pratiques des objets numériques. Dans ce contexte, les Humanités Numériques forment l'ensemble des savoirs et des recherches portant sur les usages et l'impact des technologies numériques sur la société et l'individu. Elles doivent donc articuler des connaissances scientifiques, des savoir-faire technologiques, mais elle a aussi une visée grand public. Former aux Humanités Numériques

est donc un projet démocratique qui vise à repérer ce qui peut être conservé de l'humanisme classique (Doueihi, 2011).

Or la culture numérique transmise en tant que telle revêt *a priori* peu d'intérêt pour l'éducation : à ce jour nous n'avons pas de preuves scientifiques qu'elle peut élever le niveau scolaire des élèves en général. L'innovation numérique ne démontre pas son attractivité au niveau des apprentissages, à part pour favoriser la répétition, *via* les flashcards. L'engagement récent des sciences cognitives pour évaluer les outils numériques est prometteur, mais la recherche dans ce domaine reste encore embryonnaire. Les investissements massifs dans la formation (plans tablette à l'école, etc.) ont aussi montré leurs limites. Dès lors, pourquoi former à la culture numérique ? Sans doute est-ce devenu nécessaire, moins pour apprendre que pour mieux insérer l'élève dans la société contemporaine, pour mieux le stimuler, jouer sur sa motivation et l'ouvrir à de nouvelles expériences.

L'enseignement de la culture numérique est pris partiellement en charge *via* des dispositifs comme le Permis Internet pour les enfants, dont se charge la Gendarmerie nationale, la Police nationale, la Préfecture de Police voire l'association Axa Prévention, qui unissent leurs forces et leurs expertises en matière de protection et de prévention contre les risques sur Internet. Cependant, est-ce suffisant ? En effet, il faut en premier lieu distinguer la formation aux usages du numérique et la formation aux Humanités Numériques et scientifiques. Cette distinction permet d'identifier un hiatus entre une approche qui viserait à fournir des savoir-faire, comme la maîtrise de la bureautique, des comportements comme la prudence sur les réseaux sociaux, et des compétences réflexives et critiques au fondement de la culture numérique et de son développement. Ainsi un des grands enjeux de l'enseignement de la culture numérique est de renforcer chez les apprenants des *compétences transversales*, en développant l'autonomie, la créativité ou le goût d'apprendre. De là, il devient possible de repenser les différentes disciplines scolaires. La révolution numérique est un donc un tournant, et non une *tabula rasa* qui laisserait oublier notre culture humaniste classique, comme le prônent les tenants d'un post-humanisme. Nous ne pouvons faire l'impasse sur notre histoire. Chaque nouveau support d'écriture, de l'imprimerie au numérique, enrichit notre pensée en termes de représentations, de méthodes et de modes d'action (Goody, 1979). Ainsi, le principe de neutralité du Net ne doit pas masquer l'absence de neutralité des interfaces créatives des grandes industries du

numérique (GAFA, Adobe…), que doivent interroger les Humanités Numériques.

Finalement, l'enseignement des Humanités Numériques gagnera à s'appuyer sur un espace pluridisciplinaire dans lequel doit pouvoir se déployer une véritable science du design des ressources, des services et des données, avec le concours de l'innovation sociale, technologique et scientifique. Pour cela sont convoquées toutes les disciplines connexes : informatique, sciences cognitives, sciences de l'information et de la communication, sciences de l'éducation, et toutes les sciences humaines et sociales.

Les offres de formation dans le supérieur

Dans cette section, nous présentons quelques offres de masters « Humanités Numériques » dans les UFR[2], les INSPE[3] et les URFIST[4]. Les cursus des UFR s'adressent à des étudiants de licence en sciences humaines et sociales. Nous poursuivons par les formations en INSPE, dont la dynamique disciplinaire est au cœur des compétences visées. Enfin, nous terminons par la formation continue des URFIST, qui s'adresse aux formateurs en activité, et où les sciences de l'information et de la communication tiennent un rôle central, contrairement aux deux catégories précédentes.

Les offres de master « Humanités Numériques » dans les UFR

Elles visent à donner aux étudiants les compétences nécessaires pour accompagner le changement en lien avec l'innovation et les transformations numériques des organisations. Selon l'ONISEP[5], il existe en 2020 plus d'une quinzaine de masters portant l'intitulé « Humanités Numériques », chacun pouvant proposer plusieurs parcours. Les étudiants intéressés par ces parcours, relativement récents (2015) mais en plein essor, proviennent principalement d'une licence SHS, mais ils peuvent aussi venir, par exemple, d'une licence en informatique. Ils visent à « exploiter l'expertise en informatique appliquée aux sciences humaines à des métiers émergents »[6]. Les

2. Unité de formation et de recherche.
3. Instituts Nationaux Supérieurs du Professorat et de l'Éducation
4. Unité régionale de formation à l'information scientifique et technique.
5. Voir http://www.onisep.fr/Ressources/Univers-Formation/Formations/Post-bac/Master-humanites- numeriques, consulté le 24/04/2019.
6. Voir https://www.lemonde.fr/campus/article/2019/01/30/a-quoi-menent-les-masters-en-humanites-numeriques_5416520_4401467.html

compétences puisent dans les dimensions juridiques et sociologiques du numérique, aussi bien que dans l'informatique, la statistique, la collecte, l'analyse et la valorisation des données, ou encore la gestion de projets numériques. Ces masters visent des carrières de cadre, de chargé de développement numérique et culturel ou de responsable de communication numérique ; plus rarement de « data scientist », « web designer » ou chargé d'études statistiques. Dans certains cas, ils débouchent sur un doctorat en sciences humaines ou des concours dans l'enseignement ou l'administration publique. Le master de l'université de Bordeaux-Montaigne indique que 40 % des étudiants sont issus d'un bac littéraire et 55 % sont des femmes.

Ces masters sont portés par des UFR en SHS (exception faite de l'UFR de mathématiques de Saint-Denis en partenariat avec un département de Sic), mais aussi par des institutions comme le CNAM, l'ENS, l'ENSSIB de Lyon, ou encore l'École Nationale des Chartes. Certains masters sont ainsi portés en cotutelle, voire en double diplôme. Un master de l'Université Paris 8 indique des débouchés qui visent explicitement le métier de formateur en entreprise[7]. De même, le master de l'Université de Bordeaux-Montaigne, avec ses trois parcours[8], permet d'accéder au métier de « responsable de formation numérique au sein d'une collectivité et/ ou entreprise » ou à celui de « responsable de FabLab ». Ce master, rattaché à l'UFR Sciences de Territoires et de la Communication et au laboratoire MICA, propose une poursuite d'étude en doctorat de sciences de l'information et de la communication. La plaquette de formation mentionne d'ailleurs des enseignements spécifiques aux sciences de l'information et de la communication orientés vers la construction du sens, la médiation, la culture numérique, la communication des organisations ou le traitement de l'information. D'une façon générale, le versant informatique de la collecte et du traitement des données est privilégié dans ces masters. C'est le cas du master du CNAM Pays-de-la-Loire « Mégadonnées et analyse sociale ». Quant au master de l'Université Paul Valéry Montpellier 3 « Médiation numérique et ingénierie

7. Il s'agit du master de Paris 8 : « master Humanités Numériques, numérique : enjeux et technologies ». Parmi les débouchés possibles, on trouve le domaine « Formation et assistance (gestion de fablab, formation TIC, conseil Saas, animation de communautés, accessibilité numérique, travail collaboratif, réseaux sociaux d'entreprise, curation web, e-réputation, etc. ». La plaquette présente, entre autres, des compétences opérationnelles, telles que l'analyse de tests, enquêtes ou observations sur un produit, entre sciences de gestion et sciences de l'information et de la communication.

8. Il s'agit des parcours « conception de projet numérique et multimédia », « dispositifs numériques éducatifs » et « document numérique et humanités digitales ».

pédagogique », qui existe au sein de l'Institut des Technosciences de l'Information et de la Communication, il est rattaché au département « sciences du langage ». Les conditions d'admission n'indiquent pas si les licences en information-communication sont privilégiées, bien que des savoir-faire et des compétences centrées sur l'analyse d'artefacts et de leurs usages pourraient être un prérequis valorisé.

Enfin, plusieurs masters envisagent l'enseignement supérieur comme un débouché. Il s'agit alors d'apporter des méthodes et des outils profitables à la recherche (collecte, traitement, visualisation des données…) plutôt que d'acquérir des compétences utiles aux missions d'enseignement[9].

Au vu de cette première analyse exploratoire et non exhaustive, la question des enjeux des Humanités Numériques en contexte de formation de formateurs est peu interrogée dans les documents institutionnels de présentation des masters que nous avons analysés. Ces derniers visent à transmettre des compétences et savoir-faire liés à la collecte, au traitement et à la visualisation des données, mais semblent moins questionner le cadre éducatif de la transmission. En effet, à la différence des masters MEEF[10] par exemple, ces *cursus* ne préparent pas aux concours d'enseignement, car cette mission est plutôt du ressort des INSPE.

Rôle des Humanités Numériques dans les offres de formation de formateur et dans les INSPE

Les formations des INSPE préparent à des concours nationaux de professeurs des écoles. Les outils numériques sont abordés du point de vue de la pédagogie, et leur appréhension ne constitue pas une matière en soi.

Les publics des INSPE sont très diversifiés : ils proviennent des lettres, des sciences humaines et sociales, de l'économie et de la gestion, des sciences et techniques. Les masters métiers de l'enseignement, de l'éducation et de la formation (MEEF), créés en 2013, sont proposés principalement aux étudiants de licence qui se destinent à ces

9. Tel est le cas du master de l'Université Paris Sciences et Lettres (PSL) « master Humanités Numériques » : « Les débouchés sont ceux de la recherche et de l'enseignement supérieur (chercheurs, ingénieur de recherche) et de l'expertise de données des sciences humaines et sociales ». Le tronc commun accorde une place importante au traitement des données : « algorithmique et programmation pour les humanités, modélisation et structuration des données, exploitation et fouille de données, méthodes quantitatives et modélisation mathématiques ».

10. Métiers de l'Enseignement de l'Éducation et de la Formation.

métiers. C'est au cours de leur formation que les étudiants de ces masters présentent, à la fin de la première année, des concours liés à la pédagogie (CRPE, CAPES/CAPET, CPE, etc.). Il est aussi possible de suivre un parcours spécifique pour accéder aux métiers de l'éducation et de la formation hors concours. À la différence des Master HN, ces enseignements portent sur la pédagogie (gestes professionnels, conduite de classe, prévention des violences scolaires, apprentissage, orientation), et ils promeuvent des thèmes structurés autour des valeurs de la République : laïcité, discriminations, citoyenneté, environnement, développement durable[11].

Malgré les déclinaisons locales des masters MEEF, les programmes sont contraints par les exigences nationales des concours et de l'organisation de formations autour de quatre à cinq mentions : premier degré (professeur des écoles), second degré (professeur des collèges et des lycées), encadrement éducatif (CPE et cadre dans l'éducation), ingénierie pédagogique et publics spécifiques. Au-delà de cette organisation, nous avons repéré trois types de modules répondant aux exigences d'une formation de formateurs aux technologies numériques[12], en lien avec l'apprentissage du code, le recours à des assistants pédagogiques ou dans le but d'innover sur le plan pédagogique.

Le premier type est incarné par les INSPE de Nantes ou de La Martinique, qui ont choisi de former au code informatique. Cette approche permet de sensibiliser les formateurs sans en faire nécessairement des spécialistes : ils sauront rédiger le code d'un lien hypertexte, mais cela ne signifie pas qu'ils en maîtrisent la sémantique et la rhétorique avancée (Bouchardon, 2014). À l'INSPE de Nantes, la mention Premier Degré du master MEEF initie au codage informatique à travers des activités de programmation ou de robotique pédagogique, dans un esprit constructiviste semblable à celui qui animait Seymour Papert qui diffusa le langage de programmation Logo. De même, mais de manière plus approfondie, le parcours « Mathématiques et Informatique » de la mention second degré du master MEEF de Nantes propose ce type d'initiation.

À l'aide d'un second type de modules, les INSPE de Poitiers, Paris, Val-de-Loire et Aix-Marseille favorisent la pédagogie par les usages des technologies numériques. Des outils et des médias numériques sont

11. http://www.devenirenseignant.gouv.fr/cid98901/de-la-licence-au-master-meef.html.
12. http://www.reseau-espe.fr/formations-et-diplomes/se-former-par-le-numerique, consulté le 28 avril 2019.

intégrés à l'apprentissage comme autant d'assistants qui renforcent l'efficacité des dispositifs de formation existants (Mœglin, 2017). Comme l'indique, par exemple, le master MEEF de premier degré de Poitiers : « *un enseignement conséquent au et par le numérique est dispensé sur la compréhension des enjeux, sur la maîtrise des technologies à but éducatif, sur la promotion des usages raisonnés dans le but de faciliter les apprentissages et la compréhension* »[13].

Enfin, un troisième type de modules permet aux INSPE de Créteil et de Strasbourg de développer l'innovation pédagogique par les usages numériques. À l'INSPE de Bretagne, une Unité d'Enseignement « Cultures numériques » vise à acculturer les enseignants au numérique (Durampart, 2016). L'INSPE de Créteil propose, quant à lui, de concevoir des situations d'apprentissage en s'appuyant notamment sur des outils (tablettes, TNI…) et des ressources numériques (banques de ressources, application de conférences, etc.).

Nous avons observé que, en dehors du Master MEEF Documentation, les enseignements sont rarement attribués aux enseignants-chercheurs en sciences de l'information et de la communication. Quant aux mentions « pratiques et ingénierie de la Formation » de ces masters, ils traitent de problématiques variées qui vont de la prise en charge de publics spécifiques jusqu'au traitement de situations d'enseignement à l'international.

Dans aucune de ces formations, l'apprenant n'envisage les technologies numériques comme formant une matière à enseigner à part entière : les technologies sont abordées comme instrument et non pas comme objet d'étude interdisciplinaire. Or, le dépassement et le décloisonnement disciplinaire sont des enjeux des Humanités Numériques qui permettent de renouer avec la figure d'un formateur humain conservant une distance critique. Ainsi les Humanités Numériques ne semblent pas, à ce jour, être au cœur des préoccupations des formations en INSPE.

Les Humanités Numériques au sein des URFIST

Les sept Unités Régionales de Formation à l'Information Scientifique et Technique (URFIST) sont des services inter-académiques de

13. http://formations.univ-poitiers.fr/fr/index/master-XB/master-XB/master-metiers-de-l-enseignement-de-l-education-et-de-la-formation-meef-1er-degre-program-master-metiers-de-l-enseignement-de-l-education-et-de-la-formation-meef-1er-degre-2/master-metiers-de-l-enseignement-de-l-education-et-de-la-formation-meef-1er-degre-site-de-poitiers-subprogram-master-meef-premier-degre-site-de-poitiers-fr.html, consulté le 24 avril 2019.

formation, de veille et d'expertise. Ils ont été créés en 1982 sous l'impulsion de la direction des bibliothèques, des musées et de l'information scientifique et technique (DBMIST) du ministère de l'éducation nationale[14]. Les URFIST visent à transmettre des compétences en matière d'usage et de maîtrise de l'information scientifique et technique (IST). Le réseau des URFIST s'est constitué en un Groupement d'Intérêt Scientifique (GIS) en 2017. Chaque URFIST est dirigé à la fois par un enseignant-chercheur en sciences de l'information et de la communication et par un conservateur des bibliothèques. Les URFIST ont sans doute été à l'avant-garde de la création de formations aux Humanités Numériques. À titre d'exemple, on peut citer des enquêtes sur les pratiques informationnelles des chercheurs (Gallezot, 2011), la diffusion de formations mais aussi de ressources sur Internet concernant les Humanités Numériques.

Les URFIST affirment explicitement sensibiliser aux Humanités Numériques[15]. Les textes de présentation de trois stages se déroulant à Rennes (le 11 mars et le 7 mai 2019) et à Strasbourg (le 7 juin 2019) évoquent un retour historique sur le concept, sur les difficultés de s'accorder sur une définition, sur ses méthodes et ses outils. Comme en témoignent les fiches de formation que nous avons lues, ces trois stages d'une journée sont bâtis sur le même modèle et adoptent les mêmes programmes et objectifs. Cependant nous n'avons pas trouvé d'autres stages proposés par les URFIST portant sur l'expression « Humanités Numériques ».

Le programme de ces journées correspond par ailleurs au sommaire d'un rapport commandité par l'Institut Français sur les Humanités Numériques et co-rédigé par Pierre Mounier[16], l'un des formateurs au côté de Marie-Laure Malingre et Annaïg Mahé[17]. À l'image des masters recensés précédemment, les URFIST se préoccupent principalement d'apporter des outils utiles aux activités de recherche, tant pour l'accès aux sources que pour l'usage de logiciels performants. La formation invite les chercheurs à s'inscrire dans la transversalité et la transdisciplinarité, à adhérer au mouvement de l'Open Research et à se structurer en réseau. Cependant, si la formation paraît très complète sur le sujet, elle n'aborde pas les questions propres aux

14. www.enssib.fr/services-et-ressources/questions-reponses/fin-de-la-dbmist.
15. Les formations sont recensées à un niveau national sur la plateforme https://sygefor.reseau-urfist.fr/#/, consulté le 24/04/2019.
16. http://pierremounier.net/cvresume/, consulté le 24/04/2019.
17. http://urfist.chartes.psl.eu/qui-sommes-nous/equipe/annaig-mahe, consulté le 24/04/2019.

grands enjeux de la formation à l'écriture numérique évoqués en première section.

Or bien des questions se posent concernant l'enseignement des Humanités Numériques, qui va au-delà de la connaissance des outils : qu'en est-il de l'enseignement de la culture numérique et de l'adaptativité des publics ? Quelle posture les formateurs doivent-ils adopter face à ces outils que connaissent peut-être déjà une partie des étudiants ? Quels usages pédagogiques les formateurs peuvent-ils envisager ? À quelles problématiques le formateur peut-il se retrouver confronté face au déferlement des ressources ? Comment traiter l'interdisciplinarité ou la transversalité en contexte pédagogique ? À l'image de la figure du pédagogue de la Renaissance, que veut-dire enseigner en contexte d'Humanités Numériques ?

Si ces aspects pédagogiques peuvent certainement être abordés lors d'autres sessions de stages, ils ne le sont pas dans ces formations courtes et denses d'une journée consacrée spécifiquement aux Humanités Numériques. Sans doute notre compréhension du programme de formation des URFIST est lacunaire : un prolongement de cette étude exploratoire pourrait consister à interroger les formateurs lors d'entretiens approfondis.

Quels rôles donner aux formateurs ?

Suite à notre étude exploratoire, la question de la formation de formateurs aux Humanités Numériques semble constituer un point aveugle. De façon plus générale, celles de la définition d'un curriculum général ou d'un référentiel de compétences en Humanités Numériques, et celles de la définition d'une pédagogie spécifique ou d'un cadre de référence ne semblent pas avoir été étudiées. Or, l'un des enjeux de la formation de formateurs est de développer un esprit analytique et critique. Comme l'écrit Françoise Paquienséguy (2017), il s'agit notamment de transmettre aux apprenants la compréhension du paradigme « technologie-société-communication ». En ce sens, le rôle du formateur doit pouvoir consister à dépasser la connaissance du simple appareillage technologique des activités humaines, pour permettre à tout un chacun de penser de nouvelles relations de travail, de nouveaux modes de production et de commercialisation de produits et de services multimédias. En tant que discipline frontière, les Sic sont à même de faire dialoguer les dimensions sociales, économiques et culturelles de l'expérience des outils numériques, tant au niveau macro (l'organisation de la société dans son ensemble),

que méso (l'organisation d'activités dans des collectifs) et micro (l'organisation de modes de pensée et d'agir individuels).

Les Humanités Numériques ne sont donc pas réductibles à un ensemble morcelé d'outils hétéroclites : elles désignent l'appropriation d'un écosystème complexe en construction, dont il est nécessaire de déchiffrer le sens et le devenir. Cette approche holistique doit penser ensemble le « faire » (l'apprentissage et la maîtrise de ces outils) et le « comprendre » (savoir porter un regard distancié sur la transformation numérique du travail et des sociétés). Les apprenants doivent pouvoir donner du sens aux données qu'ils collectent, traitent et diffusent. De fait les formations en Sic n'ont pas pour vocation de former des « codeurs » (à l'image de l'École 42, maintes fois citée pour ses pédagogies supposées innovantes), mais plutôt des « décodeurs » des phénomènes sociaux et techniques. Il s'agit ainsi de confirmer le rôle des Sic en tant que *sciences de l'interprétation*.

Par ailleurs, tout ceci ne peut se faire sans développer une réflexion conjointe avec les milieux professionnels au sujet des enjeux sociotechniques de la transformation numérique des organisations, sachant que celles-ci avancent souvent « à l'aveugle » dans leurs stratégies, et qu'elles sont aussi en demande de partenariats renforcés avec les institutions académiques d'enseignement et de recherche. Dès lors, le rôle des chercheurs en Sic pourrait être de former les formateurs à l'apprentissage des compétences numériques en situation : lutte contre les *fake news* dans l'espace public, diffusion d'une culture technique et critique dans les champs de l'éducation et de la formation, et transformation numérique des organisations. C'est en effet par leurs connaissances des champs professionnels et par leurs pratiques singulières que les Sic peuvent articuler les enjeux que nous avons soulevés. En tant que discipline à la frontière d'autres savoirs disciplinaires (sciences de l'éducation, sociologie, économie, linguistique…), les Sic ont la capacité de proposer des formations qui font appel aux Humanités Numériques comme ensemble de savoirs et de savoir-faire critiques et qui mettent les étudiants en situation de conception de produits et de services innovants en réponse à de grands enjeux sociétaux et aux controverses actuelles portées par des partenaires commanditaires (entreprises, collectivités locales…). Ces enjeux et controverses concernent des thèmes plus larges comme le réchauffement climatique, la démocratie participative, etc. Cours théoriques et cours en travaux dirigés pourraient alors s'articuler à ces projets *a priori* collectifs, tout en stimulant les étudiants pour leur permettre de jouer un rôle dans la production de contenus, mais

également de dialoguer entre eux sur ces questions. Ainsi habiletés numériques, esprit critique et démarches de créativité pourraient se développer mutuellement.

Compétences et savoir-faire info-communicationnels et réflexifs

Confronté aux enjeux des Humanités Numériques, le formateur sera amené à mobiliser des compétences informationnelles et réflexives renforcées. En termes de formation, les implications sont doubles. D'une part, il s'agit de former aux capacités de chercher, de localiser, d'évaluer, d'utiliser une information et de créer de la valeur à partir d'elle. D'autre part, il s'agit de développer chez l'apprenant une réflexion critique et argumentée. Cette prise de distance critique et réflexive doit se faire, non seulement par rapport à l'information elle-même, mais par rapport aux comportements et aux usages informationnels (Serres, 2007). Ces comportements et usages impliquent le contrôle de soi et de ses productions, ainsi que la bienveillance avec les autres (Le Deuff, 2011). Ils se révéleront comme étant des moyens d'action sur le monde, autrement dit comme un préalable à la participation effective des formateurs et des formés à la société de l'information (Gardies, 2011, p. 201), afin d'aborder les industries créatives de telle sorte que chaque individu puisse être un entrepreneur, et que l'usager-consommateur devienne un auto-producteur culturel (Bouquillion, Miège, Moeglin, 2015).

Deux grands pôles de compétences et de savoir-faire, dont toute la difficulté est de les mettre en interaction, sont à développer dans la formation de formateurs pour tenir ensemble ces enjeux quel que soit le secteur professionnel :

- L'apprentissage des outils de traitement numérique de l'information, que ce soit en production multimédia (outils de création de sites web, d'images fixes ou en mouvements…) ou en analyse des données, avec ou sans formation à la programmation informatique. Les approches sémio-pragmatiques du numérique, de création et design multimédia et transmédia, comme celles sur les documents et le traitement des données pourraient alors être mobilisées.

- La capacité à comprendre les enjeux socio-économiques de la numérisation de la société et l'approche critique de la création à trois niveaux (les activités intellectuelles, les modes d'organisation du travail et l'organisation industrielle et marchande). L'approche critique des

industries culturelles, de leurs modèles socio-économiques et celle de la communication des organisations pourraient apporter leurs acquis.

En effet, si les mathématiques, voire les sciences de gestion, sont bien placées dans le pôle de l'analyse statistique des données, l'informatique dans celui de la production de systèmes d'information et les sciences économiques dans celui de la modélisation de l'écosystème numérique, les Sic ont un rôle à jouer au niveau des interactions à mettre en place entre les deux pôles : « Humanités » et « Numérique ». De nombreux modèles, théories ou concepts issus des Sic sont en mesure d'investir le champ des Humanités Numériques et la formation de formateurs. Ces modèles sont à extraire pour une part du domaine de l'information-documentation bien sûr, notamment par certaines équipes fortement impliquées dans la définition de compétences informationnelles comme le Groupe de Recherche sur les Cultures et la Didactique de l'Information (GRCDI[18]), créé à l'initiative de l'URFIST de Rennes. On évoquera aussi la critique des industries culturelles et de leurs modèles socio-économiques, un champ largement plébiscité par les chercheurs en Sic qui font de la critique des dispositifs numériques le fer de lance de leurs recherches. Quant aux travaux sur l'industrialisation de l'éducation, ils doivent pouvoir être confrontés à la question du déploiement de l'Intelligence Artificielle ou de la formation à distance dans l'éducation. Enfin le champ émergent de la création et du design multimédia semble fructueux pour permettre aux Humanités Numériques d'investir le champ de la pratique d'écriture et de design de l'expérience interactive, champ qui recourt à la fois à des savoir-faire et à des dispositions réflexives et critiques (Hulin, 2018).

Une voie possible d'articulation réside dans la mise en place de pédagogies innovantes et de nouveaux lieux de médiation : Fablabs, ateliers créatifs, dispositifs immersifs, friches artistiques… Ces dispositifs favorisent l'apprentissage de méthodes dites « agiles » de gestion de projet en relation avec des commanditaires intervenants au sein même des formations. Ils permettent l'expression de la créativité et de la dimension participative dans l'évaluation des étudiants. Ils invitent à plus de collaboration entre enseignants, chercheurs, apprenants et professionnels dans la production d'une œuvre collective capable de prendre en compte les enjeux sociaux, économiques et culturels. C'est sur ce dernier point que devrait alors insister la formation des formateurs : former à la mise en place de

18. https://www.sites.univ-rennes2.fr/urfist/presentation_grcdi.

configurations d'apprentissages ouverts. Ces configurations pourront intégrer, sous une forme synthétique, les programmes pédagogiques d'autres disciplines : sciences fondamentales, sciences de gestion… Elles pourront aussi s'intégrer dans l'enseignement en collège et au lycée.

Les Sic apporteraient ainsi un éclairage pertinent, à l'heure où les élèves sont de plus en plus confrontés à des pratiques qui soulèvent de graves questions éthiques.

Conclusion

En tentant de dessiner les enjeux de la formation de formateurs, nous avons souligné que les Humanités Numériques se sont constituées en un courant de recherche, et non en un ensemble structuré du savoir.

Nous avons tenté, d'une part, de dresser une liste de ces questions vives et, d'autre part, de développer un point de vue critique sur l'offre de formation préalable à la formulation de propositions et de recommandations. Une des limites de ce chapitre serait de savoir si cette question de la place des Humanités Numériques dans la formation se pose de la même manière dans d'autres pays, ou bien si elle a été résolue ailleurs.

Au final, la formation de formateurs en Humanités Numériques doit développer des compétences en termes d'usages de la technique (de savoir-faire mobilisables dans des situations connues), mais aussi de diffusion d'une culture réflexive et critique pour former des acteurs et citoyens libres et responsables ; également, elle doit développer la créativité pour faire face à des situations inédites, pour répondre aux enjeux économiques et sociaux des sociétés fortement technologisées ou en phase de le devenir. Faire face à des situations inédites, c'est justement ce que permet cette culture réflexive et critique centrée sur l'humain, sachant que celui-ci se transforme au contact de la machine. Les humanistes de la Renaissance décentraient l'attention sociale et intellectuelle portée sur Dieu et ses créatures vers l'Être humain créateur. De la même façon, les humanistes du XXI^{e} siècle gagneront à décentrer l'attention portée à la machine vers l'Humain : ce dernier doit demeurer maître et possesseur de la technique. S'il accepte que la machine le transforme, c'est à partir de sa volonté éclairée et de son expérience. Or, en la matière, tout reste à construire car les formateurs aux Humanités Numériques ne sont pas encore préparés à cela. Ces formations doivent-elles prendre la forme de modules communs à toutes les formations en Sic ? Doivent-elles être portées

par les INSPE et les URFIST ? Ou, au contraire, faut-il construire des formations spécifiques de niveau master ? Selon nous, ces questions restent ouvertes.

POSITIONNEMENT

Face aux enjeux multiples de la numérisation de la société, les sciences de l'information et de la communication, discipline frontière, ont l'espace pour développer une offre de formation de formateurs reliant habiletés numériques à esprit critique et aux démarches créatives liées à la transformation numérique. Une pédagogie par projet, favorisant la création de produits et services numériques, organisant différemment cours théoriques et pratiques, serait alors à réfléchir et, peut-être, à généraliser. C'est en effet par leurs connaissances des champs professionnels et des pratiques inhérentes que les Sic pourraient articuler ces enjeux entre eux. En tant que discipline à la frontière d'autres savoirs disciplinaires (sciences de l'éducation, sociologie, économie, linguistique…), les Sic ont la capacité de proposer des formations qui font appel aux Humanités Numériques comme ensemble de savoirs et de savoir-faire critiques. Elles mettent les étudiants en situation de conception de produits et de services innovants en réponse à de grands enjeux sociétaux et controverses actuels portés par des partenaires commanditaires (entreprises, collectivités locales…). Ces enjeux concernent le réchauffement climatique, la démocratie participative, etc. Cours théoriques et cours en travaux dirigés pourraient alors s'articuler à ces projets *a priori* collectifs, qui pourraient par ailleurs avoir l'avantage de demander aux étudiants, non seulement de jouer un rôle dans la production, mais, également, de dialoguer entre eux. Ainsi habiletés numériques, esprit critique et démarches de création pourraient aller de pair.

BIBLIOGRAPHIE

Bouchardon S., 2014. L'écriture numérique : objet de recherche et d'enseignement, *Les Cahiers de la SFSIC*, juin 2014, pp. 225-235.

Bouquillion P., Miège B., & Moeglin P., 2015. Industries du contenu et industries de la communication. Contribution à une déconstruction de la notion de créativité. *Les Enjeux de l'Information et de la Communication*, 16(3B) [en ligne]. Consultable à l'adresse : http://lesenjeux.u-grenoble3.fr/2015-supplementB/01-Bouquillon-Miege-Moeglin/index.html. Dernière consultation le 12/03/2020.

Dacos M., Mounier P., 2014. *Humanités Numériques : état des lieux et positionnement de la recherche française dans le contexte international*. Institut français. Consultable sur https://hal.archives-ouvertes.fr/hal-01228945. Dernière consultation le 12/03/2020.

Deuze M., 2006. « Participation, remediation, bricolage : Considering principal components of a digital culture ». *The information society*, *22*(2), pp. 63-75.

Doueihi M., 2011. *Pour un humanisme numérique*. Paris, Le Seuil, 177 p.

Durampart M., 2016. « La forme scolaire en action traversée par l'école numérique ». *Revue française des sciences de l'information et de la communication*, (9). En ligne : https://doi.org/10.4000/rfsic.2492. Consulté le 28 avril 2019.

Gallezot G., 2011. « Enquête sur les pratiques informationnelles des chercheurs : quelques résultats ». *URFIST Info*, 4 mars 2011. En ligne : http://urfistinfo.hypotheses.org/1901. Consulté le 11 septembre 2019.

Gardies C. (dir.), 2011. *Approche de l'information-documentation : concepts fondateurs*. Toulouse, Cépadues.

Goody J. (1979). *La raison graphique*. Les Éditions de Minuit.

Hulin T. (2018). How critical thinking should be at the heart of digital interactions? ESSACHESS – *Journal for Communication Studies*, 11(2(22)), pp. 85-105.

Le Deuff O. (2011). *La formation aux cultures numériques*, Limoges, FYP Éditions, 2012

Manovich L. (2002). *The language of new media*. The MIT press.

Mœglin P. (dir.) (2016). *Industrialiser l'Éducation : anthologie commentée (1913-2012).* Université Paris 8, Saint-Denis, Presses universitaires de Vincennes.

Mœglin P. (2017). Du manuel scolaire aux systèmes numérisés d'information et de communication : le cas de l'école primaire. *Hermès, La Revue*, *78*(2), pp. 65-71.

Mounier P. et Mahé A. (2013). Digital Humanities–Humanités Numériques, présentation à l'Urfist de Paris, 28 juin et 3 octobre 2013.

Paquienséguy F. (2017). Manifeste pour un positionnement des Sciences de l'Information Communication (Sic) vis-à-vis des *Digital Studies* (DS) et autres mutations du Numérique », *Revue française des sciences de l'information et de la communication* [En ligne], 10 | 2017, mis en ligne le 01 janvier 2017, consulté le 28 avril 2019. En ligne : http://journals.openedition.org/rfsic/2630.

Serres A. (2007). Questions autour de la culture informationnelle in *The Canadian Journal of Information and library science*, vol. 31, n° 1, pp. 69-85.

Serres M. (2011). *Petite Poucette. Les nouveaux défis de l'éducation*. Discours de Michel Serres de l'Académie Française, Séance solennelle du 1er mars 2011. En ligne : http://www.academie-francaise.fr/petite-poucette-les-nouveaux-defis-de-leducation.

CHAPITRE XIV

Le défi de la formation aux environnements numériques
Vers un socle de compétences

Arnaud Mercier
Université Paris 2, CARISM
Nicolas Pélissier
Université Côte d'Azur, Sic.Lab
Éric Triquet
Avignon Université, Centre Norbert Elias

Résumé : Le développement des technologies numériques et les innovations qui les soutiennent ont des répercussions fortes sur le plan sociétal. L'appropriation de ces technologies génère notamment des dérives qui fragilisent le fonctionnement de l'espace public démocratique. Ce constat rend indispensable à la fois un travail de sensibilisation à ces dérives et une formation aux outils numériques pour l'ensemble des citoyens. Ce chapitre identifie les points d'appui pour une telle formation. Il insiste sur la nécessité de dépasser une approche instrumentale, déterministe et techniciste, au profit d'un socle commun de connaissances et de compétences. Il envisage l'apprentissage de la citoyenneté numérique, en suivant une gradation des compétences et du niveau de prise de conscience des enjeux. Il s'agit de former des acteurs utilisant au mieux les performances du monde numérique dans leur vie sociale et professionnelle, mais avec la pleine conscience de leurs responsabilités dans l'écosystème relationnel du web. Les auteurs proposent donc de faire de la notion de citoyenneté numérique un concept opératoire dans la formation aux Humanités Numériques. À l'université, comme en formation continue, cet enseignement pourra s'appuyer sur quatre piliers que sont l'histoire et les enjeux du numérique, les principes du codage et de l'algorithmie, les techniques d'écritures numériques et les méthodes de veille, de traitement et d'interprétation des informations.

Mots-clés : littéracie numérique, citoyenneté numérique, compétences, algorithmie, écritures numériques, grammaire communicationnelle, curation d'informations.

Abstract: The development of digital technologies and the innovations that support

them have strong societal repercussions. The appropriation of these technologies generates in particular drifts which weaken the functioning of the democratic public space. This observation makes essential both awareness-raising work on these drifts and training in digital tools, for all citizens. This chapter identifies the points of support for such training. He insists on the need to go beyond an instrumental, deterministic and technical approach, in favor of a common base of knowledge and skills. It envisages learning digital citizenship, following a gradation of skills and the level of awareness of the issues. The objective is to train social actors so that they make the best use of the performance of the digital world in their social and professional life, but with full awareness of their responsibilities in the relational ecosystem of the web. The authors therefore consider the notion of digital citizenship as an operational concept for training Digital Humanities. At university, as in continuing education, this training must be based on four pillars : the history and challenges of digital technology, the principles of coding and algorithmics, digital writing techniques and methods of curation, processing and interpretation of information.

Keywords: digital literacy, digital citizenship, skills, algorithm, digital writing, communication grammar, news curation.

Elle semble si loin, l'époque pionnière de l'internet où l'utopie technologique laissait paraître un optimisme sans borne pour faire des outils numériques en ligne un ferment de démocratie et de progrès social. Même l'un de ses promoteurs emblématiques, Tim Berners Lee, en appelle aujourd'hui à sa refondation pour retrouver l'esprit des origines, loin de la marchandisation, de la publicité ciblée, du contrôle social et des manipulations de l'information. L'appropriation des technologies numériques a ouvert la voie à de nombreuses dérives qui fragilisent le fonctionnement de l'espace public démocratique. Cela rend d'autant plus indispensable un travail de formation et de sensibilisation des citoyens aux outils numériques pour l'ensemble de la société. Nous distinguerons ici trois types de dérives à caractère pathologique.

La première catégorie est celle des *fake news*, catalysées par les réseaux socionumériques. Certes, ce problème sensible et controversé n'est que le « nouveau visage d'un vieux problème » (Sauvageau,Thibault et Trudel, 2019, p. 21), en d'autres temps identifié par les termes de rumeur, propagande ou désinformation. Il n'empêche, le numérique donne à ce « vieux problème » une caisse de résonnance amplifiée : « les fausses nouvelles ne sont plus seulement l'apanage de la presse avec ses canulars et propagandes à la solde d'États ou autres groupes. La popularité grandissante d'Internet change la donne et on assiste à une popularisation des procédés de désinformation. Désormais, l'Internet est aussi un espace convoité par des internautes sans motivations politiques qui cherchent à tirer profit de ce marché uniquement pour faire des sous » (Sauvageau,Thibault et Trudel *op. cit*., p. 6). Au-delà

de la seule question des nouvelles erronées, ce ne sont pas seulement les champs du journalisme et de la communication politique qui sont bouleversés, mais l'ensemble de nos pratiques sociales désormais baignées dans un climat moral que Gilles Gauthier qualifie de « post-factualiste » et dont nous portons chacun d'entre nous une part de responsabilité, comme producteurs, médiateurs, promoteurs ou simples récepteurs plus ou moins consentants. Gauthier y voit un « renoncement à l'idéal d'exactitude au profit d'un idéal de sincérité » (Gauthier, 2018) qui transforme en profondeur notre rapport ontologique à la vérité et la réalité.

Une deuxième catégorie de dérives est celle qui concerne les incidences problématiques du numérique sur les comportements de l'individu en société et au travail. Parmi ces dérives figurent les tendances addictives relevant de différentes formes de cyberdépendance altérant la vie sociale (Lardellier et Moatti, 2014), mais aussi des manifestations encore plus inquiétantes relevant du cyberharcèlement et susceptibles de menacer la vie et la santé de ses victimes, au demeurant de plus en plus nombreuses : près d'un collégien sur cinq serait concerné. Les pratiques « d'ensauvagement du web » (Mercier, 2018) sont hélas bien réelles.

Au niveau des environnements de travail, de nombreuses recherches en Sic en communication organisationnelle (voir notamment Carayol *et al.*, 2016) mettent en évidence les conséquences parfois négatives de la transformation digitale des entreprises : accélération des tempos et pression de l'urgence sur les collaborateurs, risques accrus d'infobésité liés à la surconsommation des courriers électroniques et autres messageries, interférence démultipliée sur la capacité et sur la qualité de travail, appauvrissement des relations sociales en milieu professionnel, phénomènes de mimétisme et de conformisme favorisés par les réseaux socionumériques au détriment de l'innovation, atteintes facilitées à la réputation de l'organisation et du collectif de travail, etc., Nicolas Oliveri (2016) qualifie ces pathologies de « risques techno-sociaux ».

Enfin, une troisième catégorie est celle de l'utilisation abusive des données personnelles. L'écosystème des objets connectés (ordinateurs, téléphones, montres, assistants vocaux…) produit et collecte de façon exponentielle des données, le plus souvent à l'insu de leurs utilisateurs par les États et les entreprises. Bien qu'associé à de nouveaux services, ce phénomène engendre un sentiment d'inquiétude lié à leur exploitation à des fins commerciales ou au contrôle accru sur les

individus. Les risques d'atteinte aux libertés individuelles sont grands. Dans son rapport annuel de 2016, la CNIL met en garde contre deux types de risques : celui de l'enfermement de l'internaute dans une « bulle filtrante » de recommandations ; une atteinte potentielle aux logiques collectives sur lesquelles sont fondées nos sociétés. En effet, la personnalisation accrue de l'information a pour conséquence la disparition d'un socle minimum d'informations partagées, condition nécessaire à un authentique débat démocratique, et de fait la fragmentation de l'espace public.

Pour enrayer ces menaces, une éducation aux médias numériques, et plus largement à l'écosystème numérique, est essentielle, et ce dès le plus jeune âge. Les travaux des chercheurs en Sic identifient les points-clés sur lesquels une telle pédagogie du numérique, loin de toute vision déterministe et techniciste, peut s'appuyer. Les participants au symposium SFSIC/CPdirsic ont relevé la nécessité de mettre en place un socle commun de connaissances et compétences pour parvenir à cet objectif. Ce socle repose sur quatre piliers fondamentaux qui seront détaillés dans ce chapitre : histoire et enjeux du numérique ; principes et techniques des écritures et de l'édition numérique ; principes du codage et de l'algorithmie ; enfin, techniques de veille, traitement et interprétation des données.

Histoire et enjeux du numérique

Un module pédagogique consacré à l'histoire du numérique, resituée dans une histoire plus vaste des idées, des sciences et des techniques, apparaît comme une première nécessité. Cet enseignement permettrait notamment de relativiser les discours par trop euphorisants ou stigmatisants sur ladite « révolution numérique », en montrant que celle-ci s'inscrit dans un continuum beaucoup plus long, fait de ruptures et de continuités, mais aussi d'autres précédentes révolutions ou contre-révolutions, dont la digitalisation n'est que la plus récente hypostase. Comme l'a montré Bertrand Legendre (2019) pour le livre, le numérique n'est en soi ni révolutionnaire, ni réactionnaire. Il ne fait le plus souvent qu'amplifier et complexifier, au risque de l'illisibilité, d'autres évolutions sociotechniques qui ont bouleversé le monde avant son avènement.

Au-delà de cette dimension historique visant à resituer les technologies dans des tendances de longue durée, comme le préconisent Miège, Domenget et Pélissier (2017, p. 16), la transformation numérique suppose que ses usagers actuels et futurs prennent en considération

un certain nombre d'enjeux majeurs : épistémologiques, juridiques, politiques, économiques…

Les enseignements d'épistémologie gagneraient ainsi à franchir les murs des facultés de philosophie, droit ou sciences pour se généraliser à des élèves de lycée et des étudiants dans le cadre d'une formation commune à la civilisation numérique. Discipline articulée autour de la genèse et de la constitution des savoirs, l'épistémologie est bouleversée en profondeur par l'écosystème numérique, ensemble complexe et instable de « technologies intellectuelles » susceptibles de transformer nos façons de lire, de dire, d'écrire, décrire et penser le monde. Bernard Stiegler évoque à son sujet une nouvelle « organologie » des savoirs : « soit on pose que la vie de l'esprit connaît un nouveau milieu auquel elle doit s'adapter mais qui ne l'affecte pas de manière ontologique […], soit au contraire […] on pose que l'évolution numérique […] constitue un nouvel âge de l'esprit, un nouvel esprit qui serait rendu possible par cette nouvelle forme d'écriture […] qui impose de repenser l'esprit lui-même en totalité » (Stiegler, 2015, p. 14). Si l'on opte pour la deuxième piste, il devient nécessaire que les structures académiques se saisissent d'un tel *pharmakon*, à la fois poison et remède aux maux de notre société, pour en expliquer les enjeux en termes de formation des savoirs.

Or, l'apprentissage des outils numériques se fait trop souvent dans un cadre de pensée dominé par une conception par trop positiviste et techniciste des relations homme-machine, fondée sur la logique binaire de la séparation (matière/esprit, connaissance/croyance, etc.) propre au logos occidental défini par Platon, formalisé par Descartes et parachevé par les paradigmes de l'informatique et des premières générations de l'intelligence artificielle. « Penser le numérique » suppose au contraire une logique intégrative et holiste, fondée sur l'échange des savoirs et l'interaction, et non la disjonction. Comme Gilbert Simondon l'a montré avec son concept de « transduction », les objets techniques ne sont pas extérieurs à nous mais incorporés à la vie elle-même (voir le terme « digital » qui renvoie à un organe vivant, le doigt), ce qui implique l'apprentissage d'un nouvel art de vivre et de penser basé sur l'association et la circulation des connaissances, au fondement d'un savoir et d'un savoir-faire interdisciplinaires.

Les enjeux juridiques sont également à prendre en compte dans les formations. En effet les technologies numériques ne représentent pas seulement une forme nouvelle de médiation, et sont aussi des outils de désintermédiation des « sachants ». Certains, par leur profession

et leur statut social, avaient coutume de monopoliser peu ou prou l'accès à l'espace public. Ils subissent aujourd'hui une contestation de la part de ceux qui les perçoivent comme appartenant aux « élites » et une concurrence dans le commentaire politique et d'actualité, voire dans la diffusion de connaissances, parfois présentées comme « alternatives ».

Les « professionnels » de la prise de parole (politiciens, leaders syndicaux, avocats, intellectuels, journalistes…) avaient plus ou moins intégré les limites de leur art. Tel n'est pas le cas avec les internautes, jeunes ou pas, qui utilisent ces outils comme un exutoire et les vivent comme une libération des formes de domination ou de contrôle social. En conséquence, l'éducation aux médias numériques doit intégrer des éléments sur les cadres juridiques qui régissent la prise de parole dans une société démocratique, car démocratie ne signifie pas avoir droit de tout dire et écrire sans avoir d'éventuels comptes à rendre. C'est le cas de la diffamation, de l'injure publique, de l'incitation à la haine raciale, de l'antisémitisme, des propos discriminatoires, ou encore du harcèlement en ligne, seul ou en meute. La transmission de ce qu'on a appelé la « netiquette », ces règles de conduite et de politesse recommandées pour maintenir entre les internautes des liens de sociabilité même par plateformes collaboratives interposées, devrait donc être assurée en même temps que sont enseignées des compétences d'utilisateurs.

Un autre défi est d'ordre politique. Si la majorité des recherches sur les usages du numérique montrent des différences sociales qu'il peut déjà être important de corriger pour éviter que le numérique ne reproduise les inégalités sociales existantes, quelques travaux sur les « non-usagers » soulignent que la fracture sociale est encore plus béante avec ceux qui n'ont pas accès à l'internet ou se sentent totalement incompétents face à cet univers technologique. Les plus démunis socialement et culturellement, et une partie des personnes âgées, peuvent ainsi être totalement exclus du monde de l'internet et voir leur capacité citoyenne d'agir se dégrader profondément.

Emmanuel Souchier (2017) rappelle ainsi que les dispositifs techniques ne sont pas « démocratiques » en soi, qu'ils ne peuvent l'être qu'à travers la culture acquise par les usagers. Selon lui, l'enjeu principal réside dans l'acquisition d'une véritable « littératie numérique », au sens d'une culture critique donnant les moyens de comprendre les enjeux posés par les médias informatisés. C'est sur ce point que doit se focaliser d'après lui toute l'ambition de l'enseignement, de

la maternelle à l'université. De son côté, Dominique Cardon (2015) attire notre attention sur le fait que l'on ne peut réduire les médias numériques à de simples outils techniques. Ils renvoient selon lui à un véritable projet politique. Accéder aux logiques qui les sous-tendent, comprendre leurs principes de fonctionnement et surtout prendre conscience des valeurs et du type de société qu'ils soutiennent sont une condition pour que les internautes « reprennent un peu la main », comme le préconise Isabelle Falque-Pierrotin, présidente de la CNIL dans le rapport annuel 2017.

Les enjeux économiques sont tout aussi décisifs. Intérêts industriels des secteurs d'activité que les pratiques numériques impactent profondément, ou intérêts financiers des sociétés qui sont devenues leaders sur ce marché et qui usent de nombreux procédés d'influence pour le rester. L'éducation à la citoyenneté numérique gagnera à susciter une double prise de conscience économique : celle des dégâts que génèrent nos consommations en gratuité plus ou moins illégales pouvant aller jusqu'à la fragilisation de secteurs comme l'industrie musicale ou la presse ; mais aussi celle des ressorts de ce que l'on nomme désormais « l'économie de l'attention » où des opérateurs comme Facebook et YouTube adoptent des techniques visant à attirer à tout prix notre attention et à provoquer des addictions qui garantissent que nous devenions des usagers réguliers voire compulsifs.

En matière d'économie du numérique l'enjeu est d'importance pour que chacun sache que nos actions ne sont pas neutres, et que l'intérêt économique de nombreux opérateurs qui vivent de la publicité en ligne les conduit à employer des procédés d'influence psychiques qui tutoient parfois la pure manipulation. D'où la nécessité pour chacun de déployer une « écologie de l'attention » (Citton, 2014).

Principes du codage et de l'algorithmie

Un raisonnement simpliste consisterait à penser que puisque l'informatique est omniprésente, il faut former tout le monde au codage, à cette langue, avec son vocabulaire et son mode de raisonnement propre. Or cela n'est pas accessible à tous et l'on peut fort bien conduire une automobile sans être mécanicien. En revanche, il peut être utile de posséder une alphabétisation minimale pour comprendre le raisonnement des codeurs. Une initiation au codage peut être considérée comme un apport utile pour permettre à tout usager numérique d'entrer en dialogue avec eux, en étant plus lucide

sur ce qu'il est possible de faire pour éviter le risque d'une pensée magique (« on peut tout faire avec l'informatique »), et en étant moins désarmé face au vocabulaire technique mobilisé. Les sciences de l'information et de la communication gagneraient à prendre en charge un tel défi, en inscrivant l'initiation au codage dans une perspective professionnelle relationnelle et communicationnelle : non pas savoir coder mais pouvoir interagir avec des codeurs.

Au-delà du codage, il apparaît essentiel que les apprenants soient initiés aux principes de fonctionnement des algorithmes. Car ceux-ci construisent une représentation du monde. Ils imposent, selon Dominique Cardon, « une hiérarchisation des valeurs qui en vient progressivement à dessiner les cadres cognitifs et culturels de nos sociétés. » (Cardon, 2015, p. 14). C'est notamment le cas des algorithmes de recommandation utilisés pour établir des modèles prédictifs de nos comportements, de façon à anticiper nos supposés besoins ou nous orienter vers des choix jugés les plus appropriés, au risque de faire disparaître la part de sérendipité dans le choix de nos consommations.

Du point de vue pédagogique, un premier objectif pourrait être de sensibiliser les apprenants aux modalités algorithmiques permettant de garantir une plus grande diversité culturelle, informationnelle, voire politique. Une autre piste serait de les amener à réfléchir sur le design des systèmes algorithmiques de façon à réduire leur opacité de fonctionnement. La mise en place de systèmes de visualisation, préconisée par la CNIL, permettrait par exemple aux usagers des plateformes de comprendre pourquoi telles ou telles recommandations leur ont été proposées, suivant quelles logiques, à partir de quels critères. Mais plus encore, en leur offrant la possibilité de générer en retour des recommandations plus appropriées, voire de tester différentes réponses en fonction de différents paramétrages, donc de reprendre la main et de se positionner dans une posture d'acteur et non d'utilisateur passif. Il s'agirait là de promouvoir un design apte à renforcer l'autonomie et la réflexivité des usagers.

Le cas des algorithmes des *moteurs de recherche*, en particulier Google, offre aussi une illustration symptomatique des problématiques soulevées par ces systèmes de représentation du réel. Comme l'ont maintes fois souligné nombre de chercheurs en Sic (voir notamment Simonnot, 2009 ; Sire, 2016), les usages et mésusages de ces moteurs, de leurs données et métadonnées, révèlent de multiples enjeux auxquels les futurs utilisateurs sont le plus souvent insuffisamment préparés.

Contrairement aux assertions autolégitimantes d'une entreprise telle que Google News, qui revendique par son traitement automatisé des actualités un niveau d'objectivité nécessairement « supérieur » à celui des humains que sont les journalistes, les moteurs ne sont pas neutres, loin s'en faut. Pour reprendre l'expression de Paul Virilio, ils constituent au contraire des « machines de vision » qui donnent à voir un monde qu'ils contribuent en bonne partie à façonner, derrière leur apparente neutralité bienveillante matérialisée par la sobriété de leur interface.

Comme plusieurs études l'ont révélé, même des professionnels avertis comme des journalistes ont tendance à utiliser Google sans précaution, au même titre qu'ils se serviraient d'un pied de micro ou d'une caméra (Pélissier et Diallo, 2015). L'une des conséquences est que ces journalistes vont privilégier les *primary definers* que sont les sources les plus institutionnelles, favorisées de facto par un algorithme « relationnel » dont l'effet est de renforcer la notoriété et l'influence de celles et ceux qui en ont déjà. Ce phénomène de renforcement joue d'autant plus que les usagers du web se sentent souvent perdus dans l'opulence informationnelle de la Toile, et se replient volontiers sur des sources jugées par eux les plus « sûres », au détriment de sources issues du web dit « indépendant » ou « alternatif ».

D'où l'importance de former ces usagers aux fonctions avancées des moteurs de recherche et autres techniques de requêtes d'information, sans oublier les méthodes de référencement. Mais au-delà de cette formation technique, ils gagneront à être sensibilisés, comme le suggèrent les travaux de Frédéric Kaplan, à leurs enjeux socioéconomiques. Les mots et syntagmes les plus utilisés par Google et ses principaux concurrents sont source de revenus de plus en plus élevés pour ces grands moteurs, et il est important que les internautes puissent mieux maîtriser les codes et la grammaire du web. Le nouveau « capitalisme linguistique » (Kaplan, 2011) qui se met en place conduit à distinguer les « ressources primaires », issues directement des langues naturelles, des « ressources secondaires » produites et influencées par divers algorithmes (textes produits par des *spambots* ou *chatbots*, traductions automatiques, articles d'information rédigés de façon artificielle…), sachant que ces derniers se perfectionnent pour passer entre les mailles du filet tissé par Google et permettre un meilleur référencement de manière quasi automatique. Le citoyen doit pouvoir maîtriser *a minima* les flux et codes linguistiques, afin de se positionner sur le marché de la requête en acteur conscient et non en victime plus ou moins consentante.

Principes et techniques des écritures et de l'édition numérique

En matière d'écriture numérique, quatre types de compétences sont visés, renvoyant à quatre dimensions : hypertexte, navigation, visualisation et scénarisation.

La maîtrise de *l'écriture hypertextuelle* implique différentes habilités. La première consiste à mettre en place des artefacts technologiques qui, une fois activés, permettent à l'internaute de cheminer d'un texte vers un autre. Cette interface technique se double par ailleurs d'une interface sémantique : liens « internes » allant d'un texte vers un autre à l'intérieur d'un même site web ; liens « externes » proposant aux usagers de visiter d'autres sites dans le but d'approfondir leurs connaissances et d'avoir accès aux sources du texte d'origine. L'objectif est de proposer aux cyberlecteurs un traitement contextuel, référentiel et récursif des informations diffusées reposant sur une *mise en boucle* rétroactive des savoirs. Cette faculté repose à la fois sur une compétence technique (conception de « balises ») et une intelligence de la complexité qui renvoie à la dimension épistémologique évoquée plus haut. Au-delà de l'hypertexte, l'écriture numérique renvoie à une logique de production-diffusion « hypermédia » : il s'agit, non pas de superposer différents médias, mais de les intégrer dans une logique de complémentarité entre les supports (texte, son, image fixe et animée).

Une deuxième compétence à acquérir est celle de concevoir des *circuits de navigation* structurants pour l'usager. Elle suppose de s'affranchir du modèle linéaire du texte littéraire et de dépasser les contraintes unilatérales et verticales les plus usuelles des médias de masse hors-ligne. Au niveau temporel, il s'agit de s'émanciper des périodicités spécifiques à ces médias pour proposer au cyberlecteur des « rendez-vous » à des horaires adaptés à leurs besoins. Au niveau spatial, il s'agit de produire une information destinée à un public diasporique et de satisfaire la demande de communautés virtuelles déconnectées de la matérialité d'un territoire déterminé.

En outre, les contenus pédagogiques d'une formation à l'édition numérique gagneront à intégrer une véritable *compétence graphique* permettant une meilleure visualisation des données et de l'information. Dans l'univers numérique, la force des messages repose sur quelques mises en forme qu'il est bon de savoir maîtriser pour profiter de tout le potentiel communicationnel de la grammaire numérique. Parmi ces dispositifs spécifiques, on trouve des éléments graphiques qui ont l'avantage de frapper les esprits, d'attirer le

regard et l'attention, et souvent de synthétiser une idée ou une démonstration en peu d'espace et peu de texte. Ils peuvent aider à faire comprendre et passer des messages de façon plus efficace, avec des vertus pédagogiques pour peu qu'on en maîtrise les modalités. Cette forme de langage graphique fait partie désormais des outils communicationnels à intégrer. On pense notamment aux frises chronologiques (plus souvent nommées par le terme anglais *timeline*) qui permettent un récit historique enrichi avec des liens hypertextes et diverses illustrations multimédia. L'infographie fait aussi partie de ces compétences nouvelles, voire la datavisualisation, mais elle nécessite une formation bien plus approfondie, avec des statistiques en sus. Ces dispositifs numériques et encore d'autres peuvent être considérés comme un socle narratif nouveau à prendre en compte par les apprenants.

Enfin, ce dernier point amène à évoquer une quatrième compétence qui renvoie à la maîtrise des *techniques de scénarisation*, actualisées à l'ère du *transmedia storytelling*. Comme certains auteurs l'ont relevé (Pélissier et Eyries, 2016), les enseignements de narratologie sont très lacunaires dans l'enseignement secondaire et supérieur francophone. Ils se limitent trop souvent à des cours sur les fondamentaux du récit et aux commentaires de texte littéraire ou de film en classe de français ou d'audiovisuel. Dans l'enseignement supérieur, exception faite des études littéraires et des Écoles d'audiovisuel, ou du « journalisme narratif » (Grévisse, 2015) les étudiants découvrent les principes de la mise en récit dans des cours d'analyse du discours encore trop rares et proposés dans les formations en Sic et SHS souvent de façon optionnelle. Ces enseignements apparaissent éloignés d'une narratologie communicationnelle, médiatique et transmédiatique (Lacelle et Lebrun, 2017) telle qu'elle pourrait être enseignée de façon plus précoce à des élèves davantage consommateurs de médias électroniques que de textes classiques et de films d'auteur. Non que ces derniers doivent être négligés, mais ils pourraient être mieux mis en relation avec les problématiques les plus contemporaines véhiculées notamment par les séries télévisées, qui suscitent de nombreuses *fan fictions* qui consistent à impliquer le récepteur dans la corédaction des scénarios de ces séries.

Veille, traitement et usage des données

L'une des caractéristiques de l'univers numérique est de générer des quantités faramineuses de données, puisque chaque clic, chaque envoi, chaque circulation, chaque formulaire rempli laissent des

traces de notre activité. Une véritable économie de la donnée s'est donc mise en place, où des géants comme Google et Facebook occupent une place de choix, capables de dessiner des profils hyper-précis de chacun en accumulant et traitant nos traces d'usage. Mais cet exercice peut aussi devenir une activité de chacun, en considérant le monde numérique comme une gigantesque base de données où il faut apprendre à se repérer. L'internet s'enrichit de centaines de milliers de pages nouvelles chaque jour. Le mouvement de « l'open data » rendant accessible librement des données officielles gagne du terrain. Bien des informations sont accessibles pour peu qu'on sache les chercher.

Dès lors, la maîtrise des techniques de veille d'information doit faire partie du *vade mecum* de l'internaute agile. Maîtrise des opérateurs de requête sur les moteurs de recherche, activation d'outils automatisés de veille et de curation, culture du partage de sa veille au profit réciproque de tous, telles sont les compétences informationnelles indispensables pour tirer tous les bénéfices de cet univers foisonnant de connaissances qu'est l'internet. Reste aussi à acquérir de bons réflexes en matière de vérification de la fiabilité des informations ainsi recueillies. Nous savons que démocratisation de la mise en ligne d'informations rime aussi avec approximation ou pire, manipulation intentionnelle. Tout ce qui est publié en ligne ne se vaut pas, loin s'en faut ! Chacun doit donc apprendre à se livrer à un travail de recoupement des sources, de vérification de la crédibilité des lieux de publication, et à ne pas se laisser abuser par des informations sensationnelles, en déployant un esprit critique éclairé.

Lorsque se sont développées les STIC (Sciences et Techniques de l'Information et de la Communication) au début des années 2000, les Sic ont mis en évidence leur capacité à attribuer du sens aux informations traitées par les machines. C'est la raison pour laquelle elles apparaissent bien placées, comme cela a été relevé lors du symposium à l'origine du présent ouvrage, pour se positionner comme des *sciences de l'interprétation* des données digitales en plus ou moins grande quantité. Elles rejoignent ainsi la mission attribuée aux humanités digitales par des auteurs comme Milad Doueihi (2011) ou Pierre Mounier (2018).

Ce dernier insiste sur les limites des outils purement quantitatifs de traitement des données de masse tels que Google les fournit à partir des millions d'ouvrages numérisés par ses soins : « la notion d'information n'est rien sans une activité d'interprétation qui lui

donne du sens et la contextualise. De la même manière que le décodage du génome humain est vain sans l'activité herméneutique qui le relie […] par exemple au risque de développer telle ou telle maladie, il est vain d'attendre du traitement automatique de Google Ngram Viewer […] qu'il nous permette de déduire d'éventuelles Lois de la culture » (Mounier, in Stiegler, 2015, p. 106).

En vue de transmettre une telle *compétence herméneutique* aux élèves et étudiants, les formateurs gagneront à mobiliser des outils issus des études littéraires telles que les techniques du récit ou de la stylistique (métaphore, métonymie, etc.), de même que les techniques de catégorisation et de cadrage produites par les sciences sociales, sans oublier les études d'argumentation, trop souvent négligées par nos instances académiques. Cela s'avère d'autant plus nécessaire que les modes de communication sur les plateformes de réseaux sociaux favorisent l'expression spontanée et sans filtre d'états d'âmes individuels souvent peu compatibles avec les exigences de la raison argumentative indispensable à la vie démocratique.

Conclusion : encourager littératie et citoyenneté numériques

Isabelle Falque-Pierrotin, présidente de la CNIL, martèle que face à des systèmes sociotechniques de plus en plus complexes et, à l'heure de la dématérialisation généralisée, il importe en premier lieu de permettre aux personnes de ne pas « perdre la main ». Il y a là un enjeu fort d'éducation qui déborde le strict champ de l'École et de l'université. Emmanuel Souchier (2017, p. 13) critique l'adage qui veut que la technologie est – en soi – susceptible de nous amener le savoir et la démocratie. Or, comme il le souligne, il n'y a pas de dispositifs techniques de lecture et d'écriture plus difficiles d'accès que les médias informatisés. Il importe donc, selon lui, de les comprendre puis de les pratiquer en citoyens, et non en consommateurs installés dans le discours dominant.

S'interroger sur la formation au numérique oblige à dépasser une approche instrumentale telle que proposée par le référentiel de type « C2I » au profit d'une approche en termes de littéracie numérique, qui « n'est pas une catégorie technique qui décrit un niveau fonctionnel minimal de compétences technologiques, mais plutôt une vaste capacité de participer à une société qui utilise la technologie des communications numériques dans les milieux de travail, au gouvernement, en éducation, dans les domaines culturels, dans les

espaces civiques, dans les foyers et dans les loisirs. » (Hoechsmann et DeWaard, 2015). Dans cette optique, le « Référentiel de formation des élèves à la protection des données personnelles à l'attention des personnels éducatifs », conçu en 2016 par l'Autorité Française de Protection des Données et la CNIL, apparaît essentiel. Il en va de même du programme EMI « Éducation aux médias et à l'information » mis en place depuis 2013 au collège dont l'objectif est clairement « de permettre aux élèves d'exercer leur citoyenneté dans une société de l'information et de la communication, former des "cybercitoyens" actifs, éclairés et responsables de demain ».

Une formation optimale aux Humanités Numériques, dans l'enseignement supérieur ou pour des populations adultes en formation continue, gagnerait donc à s'appuyer sur quatre piliers fondamentaux que sont l'histoire et les enjeux du numérique, les principes du codage et de l'algorithmie, les techniques d'écritures numériques et les méthodes de veille, traitement et interprétation. Selon nous, les Sic ont une longueur d'avance dans ces domaines par rapport à d'autres disciplines et elles bénéficient d'une expérience conséquente permettant de les relier entre elles. Par exemple, les travaux précurseurs des années 1980-1990 sur les transformations du document, sur l'indexation ou les usages des TIC, sur les cartographies indicielles ont permis aux Sic de penser le « numérique » avant sa généralisation. Face à la montée des projets interdisciplinaires labellisés « Humanités Numériques » (dont le contenu est d'ailleurs parfois très flou…), Olivier Le Deuff (2012) propose aux chercheurs en Sic de ne pas se replier sur une posture identitaire, mais plutôt de participer plus activement au dialogue interdisciplinaire et de retravailler leurs fondements épistémologiques.

À l'issue de notre analyse, nous suggérons de faire de la notion de *citoyenneté numérique*, souvent évoquée dans la vie sociale ou dans les référentiels d'éducation aux médias pour enfants et adolescents, un véritable concept opératif dans le travail pédagogique de formation. La citoyenneté numérique représenterait un ensemble de compétences permettant d'acquérir des agilités autorisant chacun à tirer pleinement profit de toutes les potentialités du monde numérique et d'en acquérir la grammaire communicationnelle. Tout cela, en ayant acquis le regard critique suffisant pour avoir conscience de ses responsabilités individuelles dans l'écosystème relationnel du web et pour agir de façon éthique, tout en ayant la lucidité de déjouer les tentatives d'influence et de manipulation qui foisonnent, afin de

faire de l'univers numérique un moyen d'émancipation personnelle et sociale.

POSITIONNEMENT

On peut concevoir l'apprentissage de la citoyenneté numérique selon une gradation qui enrichit les compétences et le niveau de prise de conscience depuis les rudiments acquis à l'école dans le cadre de l'éducation aux médias et à l'information, où on apprend à devenir vigilant tout en étant respectueux d'autrui, jusqu'à l'enseignement pour les étudiants et les adultes où il s'agit de devenir des acteurs engagés dans la vie sociale et professionnelle utilisant au mieux les performances du monde numérique et en déjouant tous les pièges.

La forme d'opérationnalisation optimale de cette ambition serait d'arriver à créer un module dédié de citoyenneté numérique courant dans tous les cursus de l'enseignement supérieur (et pouvant être reproduit dans des parcours dédiés de formation continue pour adultes), avec pour objectifs la maîtrise d'une série d'outils numériques, la confrontation avec des situations de mésusage afin de revoir la netiquette et de sensibiliser chacun à l'esprit de responsabilité, et le dévoilement des techniques de dévoiement des usages citoyens de l'internet afin d'acquérir les bons réflexes de protection.

Les enseignants-chercheurs en Sic sont particulièrement armés pour animer ce type de formation et pour assurer, au sein des INSPE, des cycles de formation pour les enseignants du secondaire qui peuvent apporter, dans le cadre de l'EMI, la première couche de sensibilisation aux exigences de la citoyenneté numérique.

BIBLIOGRAPHIE

Carayol V. *et al*., 2016. *La laisse électronique, les cadres débordés par les TIC*, Pessac, Maison des Sciences de l'Homme d'Aquitaine, 146 p.

Cardon D., 2015. *À quoi rêvent les algorithmes ? Nos vies à l'heure des big data*, Paris, Le Seuil, 112 p.

Citton Y., 2014. *Pour une écologie de l'attention*, Paris, Seuil, 320 p.

Doueihi M., 2011. *Pour un humanisme numérique,* Paris, Le Seuil, 177 p.

Gauthier G., 2018. « Le post-factualisme : réalité, communication, information et débat public », *Communication*, 35 (1), 2018. https://journals.openedition.org/communication/7530. Dernière consultation le 12/03/2020.

Grévisse B., 2015. *Écritures journalistiques : stratégies rédactionnelles, multimédia et journalisme narratif*, Bruxelles, De Boeck, 262 p.

Hoechsmann M., DeWaard H., 2015. *Définir la politique de littératie numérique et la pratique dans le paysage de l'éducation canadienne*, Ottawa, BabiloMédias. https://habilomedias.ca/sites/mediasmarts/files/publication-report/full/definir-litteratie-numerique.pdf. Dernière consultation le 12/03/2020.

Kaplan F., 2011. « Vers le capitalisme linguistique : quand les mots valent de l'or », *Le Monde Diplomatique*, novembre, pp. 18-19.

Lacelle N., Lebrun M., 2017. « L'écriture transmédiatique dans le processus d'appropriation d'univers narratifs », in M.-J. Fourtanier, Fr. Le Goff (dir.), *Les formes plurielles des écritures de la réception*, Namur, Presses universitaires de Namur, pp. 237-258.

Lardellier P., Moatti D., 2014. *Les ados pris dans la Toile*, Paris, Le Manuscrit, 208 p.

Le Deuff O., 2012. « Littératies informationnelles, médiatiques et numériques : de la concurrence à la convergence ? », *Études de communication*, 38, pp. 131-147.

Legendre B., 2019. *Ce que le numérique fait aux livres*, Grenoble, PUG, 138 p.

Mercier A., 2018. « Comment expliquer l'ensauvagement du web », *La revue des médias*, 20 avril. https://larevuedesmedias.ina.fr/comment-expliquer-lensauvagement-du-web. Dernière consultation le 12/03/2020.

Miège B., Domenget J.-Cl., Pélissier N., (dir.), 2017. *Temps et temporalités en information-communication : des concepts aux méthodes*, Paris, L'Harmattan, 200 p.

Mounier P., 2018. *Les Humanités Numériques, une histoire critique*. Paris, Editions de la MSH, 176 p.

Oliveri N., 2016. « De l'otakisme à la cyberdépendance, une étude France-Japon », *Terminal*, n° 119. https://journals.openedition.org/terminal/1534. Dernière consultation le 12/03/2020.

Pélissier N., Eyries A., 2016. « Le journalisme narratif, enjeu de formation professionnelle et/ou d'éducation citoyenne ? », *Multimed,* 4, pp. 153-162.

Pélissier N., Diallo M., 2015. "Le journalisme à l'épreuve des dispositifs socionumériques d'information et de communication", *Revue Française des Sciences de l'Information et de la Communication*, 6, https://rfsic.revues.org/1449. Dernière consultation le 12/03/2020.

Sauvageau Fl., Thibault S., Trudel P., (dir.) 2019. *Les fausses nouvelles : nouveaux visages, nouveaux défis*, Paris, Hermann, 284 p.

Simonnot B., 2009. « Culture informationnelle, culture numérique : au-delà de l'utilitaire ». *Les Cahiers du numérique*, 5 (3), pp. 25-37.

Sire G., 2016. *Les moteurs de recherche*, Paris, La Découverte, 128 p.

Souchier E., 2017. « Et demain j'apprends quoi ? Le leurre démocratique du code informatique », *Le Français aujourd'hui*, 196 (1), pp. 11-24.

Stiegler B. (dir.), 2015. *Digital Studies*, Paris, FYP, 192 p.

CONCLUSION GÉNÉRALE

Penser, former, transformer

Nicolas Pélissier et
Françoise Paquienséguy

L'objectif essentiel de la réflexion collective sur les Humanités Numériques, entamée en 2015, était de formuler, à l'issue de nos échanges, des positions et propositions. Émises au sein de notre communauté disciplinaire, elles sont susceptibles d'être reprises par d'autres communautés et d'inspirer nos institutions d'enseignement et de recherche. À l'issue de chacun de ses chapitres, les auteurs ont proposé un encadré résumant ces positions et propositions au regard des quatre catégories d'enjeux : épistémologiques, politiques, méthodologiques et pédagogiques. Quelles sont leurs suggestions majeures ?

Agir avec six principes

À propos des enjeux théoriques et épistémologiques soulevés par le développement des Humanités Numériques, certains contributeurs incitent les chercheurs en Sic à faire preuve d'un devoir de *pratique*. Il apparaît essentiel de faire usage des outils proposés par les Humanités Numériques pour mieux évaluer les interactions entre sciences, tecHumanités Numériquesologies et sociétés. Ils évoquent aussi un impératif de *critique* : il s'agit de montrer les apports du numérique aux humanités, mais aussi d'évoquer les risques et les dysfonctionnements en prenant en compte notamment les stratégies de ses acteurs. Une autre exigence est mise en avant : la *formation*, qui vise à transmettre

et discuter les savoirs et savoir-faire liés au numérique aux futurs et actuels praticiens et chercheurs. D'autres propositions insistent sur un devoir de *mémoire* : prendre en compte l'ensemble des travaux, notamment en Sic, qui ont porté sur le traitement de l'information avant l'informatique en réseaux. Elles évoquent également un devoir d'*ouverture*. Les Sic revendiquent un savoir-faire, éprouvé de longue date, fondé sur l'articulation des recherches sur l'apport des outils numériques aux SHS avec l'étude des contextes sociaux dans lesquels s'utilisent ces outils, ainsi que des représentations sociales des données qu'ils produisent. Elles paraissent également bien équipées pour interroger le sens que ces données construisent et véhiculent, dans la perspective d'un devoir d'*interprétation.*

Penser ensemble politique et éthique

En outre, les Sic gagneront à interroger au plus près les enjeux politiques et éthiques liés aux Humanités Numériques et aux discours publics qui accompagnent leur développement. Il s'agit d'abord d'évaluer la *dimension idéologique* qui sous-tend de tels discours (politiques, médiatiques, publicitaires, académiques…) en les resituant dans une histoire renvoyant notamment à la science documentaire des années 1930, à la cybernétique des années 1940-1950, ou encore aux réflexions sur l'informatisation de la société et des établissements scolaires dans les années 1970-1980. Sans oublier les discours euphorisants sur les « Autoroutes de l'information » (années 1990), la « Société de l'information » (années 2000) ou encore la « Société Numérique » au cours de la dernière décennie. Il s'agit aussi de resituer ces discours d'accompagnement dans le cadre d'une *économie globalisée de la production scientifique* et de l'industrialisation de la logique scientifique appliquée aux SHS, dans la mesure où l'intérêt des producteurs des outils des Humanités Numériques (algorithmes, logiciels, interfaces…) ne correspond pas forcément à celui de leurs usagers, de plus en plus nombreux, diversifiés et mieux informés. D'où l'importance d'une *réflexion éthique* portant sur le numérique comme objet et outil de la recherche en SHS. Les chercheurs en Sic semblent déjà bien engagés dans cette démarche réflexive qu'ils gagneront à poursuivre à plusieurs niveaux : définition des problèmes de recherche, conformité des choix méthodologiques, production et traitement des données, diffusion des résultats. Ils prendront en compte les particularismes éthiques nationaux, tout en revendiquant une forme d'universalité permettant le développement de standards internationaux.

Renouveler les méthodologies

Les contributeurs se sont positionnés sur les questions de méthode suscitées par les Humanités Numériques. Ici encore, il apparaît important de faire fructifier la recherche en Sic engagée depuis la fondation institutionnelle de la discipline au milieu des années 1970, notamment les travaux qui ont porté sur les dispositifs de médiation, la réception des médias ou les usages des TIC. Certains des auteurs de ces travaux se sont notamment posé la question, cruciale pour les Humanités Numériques, de la *traduction des données issues des enquêtes en significations*. Dans l'écosystème numérique, cette question reste vive et les contributeurs invitent les chercheurs à faire preuve de vigilance critique dans le choix et l'application des méthodes « natives » digitales, *en évitant le double piège du tout-informatique et du tout-quantitatif*. Dans la perspective d'une science ouverte, cette vigilance doit aussi être de mise lorsqu'il s'agit de partager les données recueillies avec des partenaires extra-académiques (entreprises, agences, ONG…) sur des plateformes avant la publication définitive des résultats des enquêtes. Mais il s'agit également de considérer les *occasions de renouvellement méthodologique* offertes par les outils numériques. Les chercheurs en Sic sont ainsi invités, dans la lignée de travaux précurseurs menés dès les années 1970-1980 au sein de la discipline, à penser ensemble chiffre et lettre, computation et interprétation, information et communication, le monde des textes et des images avec celui des machines et des algorithmes. Cela semble d'autant plus important que les « sciences du numérique », telles qu'elles sont enseignées actuellement en milieu scolaire, ont tendance à privilégier l'apprentissage du code, des logiciels et des architectures (l'univers du chiffre et sa logique binaire) ; et ce, au détriment de l'étude des textes qui ont permis, organisé et parfois limité le développement de ces outils.

Former aux littératies numériques

Cette remarque, en dernier lieu, invite les chercheurs en Sic à interroger les contenus pédagogiques. D'autant que les Sic, comme l'atteste la cartographie des formations proposée dans l'ouvrage, sont très présentes dans nombre de cursus consacrés aux Humanités Numériques, ce qui constitue une particularité du système académique français par rapport à celui d'autres pays européens. Parmi les explications de ce constat, on retrouve l'articulation entre information et

communication évoquée plus haut, mais aussi celle entre la maîtrise des outils, la proximité des formateurs avec le monde professionnel et une forte tradition théorique. À ce sujet, les contributeurs évoquent la nécessité de doter les apprenants d'une triple compétence associant habiletés numériques, esprit critique et démarche créative suscitée par les demandes et besoins des partenaires institutionnels des formations. Ils suggèrent aussi de dépasser l'approche instrumentale, déterministe et tecHumanités Numériquesiciste, qui sous-tend trop souvent la formation aux Humanités Numériques, au profit de la mise en place d'un socle commun de connaissances et de compétences. Ce socle reposerait sur l'apprentissage de quatre piliers fondamentaux : l'histoire et les enjeux (épistémologiques, politiques, éthiques, juridiques, économiques…) de l'écosystème numérique ; les principes du codage et de l'algorithmie (initiation aux méthodes, compréhension du fonctionnement des algorithmes de référencement, des systèmes de recommandation ou encore des plateformes de réseaux sociaux…) ; les tecHumanités Numériquesiques d'écriture et édition numérique (compétences graphiques et design des interfaces, tecHumanités Numériquesiques narratives, maîtrise des circuits de navigation…) ; enfin, les méthodes de veille, traitement et interprétation des données (outils de curation, indexation, sémantisation, visualisation ou encore diffusion de ces données, principes de la science ouverte…). Ce socle commun pourrait constituer la base d'un apprentissage élargi d'une *citoyenneté numérique* qui pourrait être enseignée de l'école élémentaire à l'université, en passant par les collèges et lycées. Les enseignants-chercheurs en Sic seraient en mesure de jouer un rôle essentiel dans la formation de formateurs aux littératies numériques à partir d'un tel corpus de connaissances : personnels des établissements scolaires (dans le cadre notamment de l'éducation aux médias et des cursus proposés par les INSPE), mais aussi de l'enseignement supérieur au regard de la mission dévolue aux URFIST dans ce processus d'apprentissage. Enfin, ce corpus serait susceptible de servir de référentiel aux acteurs du monde professionnel et des organisations soucieuses de doter leurs collaborateurs de savoir-faire, mais aussi de savoir-être face aux problèmes contemporains posés par l'usage renforcé du numérique au travail.

La richesse et la diversité de ces propositions reflètent le pluralisme des orientations épistémologiques, courants de pensée, domaines d'investigation et pratiques méthodologiques des sciences des Sic. Celles-ci gagneront à participer activement au mouvement des Humanités Numériques, tout en produisant une pensée critique sur ses enjeux.

Elles contribueront de la sorte à transformer ce mouvement en acceptant, avec lucidité, d'être transformées en retour. Depuis près d'un demi-siècle, la recherche en Sic n'a eu de cesse de déconstruire et démystifier les discours sur les tecHumanités Numériquesologies de l'information et de la communication, tout en contribuant au développement et aux usages de leurs outils. Un paradoxe ? Plutôt une marque de fabrique, une singularité permettant de mieux penser les Humanités Numériques, former les acteurs de celles-ci et transformer leurs environnements.

Au-delà des discours prophétiques à visée performative, que recouvre le projet des Humanités Numériques ? Un domaine académique émergent, croisant l'informatique et autres sciences de l'ingénierie avec les recherches en arts, lettres, langues et sciences humaines et sociales ? Des formations et diplômes aux contenus et contours attrayants pour les étudiant·es mais aux débouchés encore incertains ? Des programmes ambitieux de recherche, à vocation interdisciplinaire, permettant un décloisonnement des savoirs et pratiques scientifiques ? Ce projet protéiforme est riche en promesses, mais l'observation attentive de sa dynamique révèle aussi certains problèmes.

Cet ouvrage questionne les Humanités Numériques à la lumière d'un processus réflexif, critique et délibératif développé au sein de la communauté des chercheurs et chercheuses en SIC. Fruit d'une coopération entre la Société Française des Sciences de l'Information et de la Communication (SFSIC) et la Conférence Permanente des Directeur·trices d'unités de recherche en Sciences de l'Information et de la Communication (CPdirsic), il s'appuie sur une réflexion collective menée depuis 2016 *via* des manifestations organisées par ces deux associations.

À la lumière de leur expérience et acquis disciplinaires, 30 auteur·trices répondent à quatre interrogations : l'écosystème numérique suscite-t-il une réelle rupture théorique et épistémologique ? Comment mieux évaluer les enjeux politiques et éthiques engendrés par la transformation numérique de nos sociétés ? Quelles sont les incidences de cette numérisation sur le travail et les méthodes des formateur·trices, tout comme des chercheurs et chercheuses ? Comment mettre en place un socle commun d'apprentissage en vue de développer de nouvelles littératies numériques ?

Coordination scientifique : Françoise Paquienséguy et Nicolas Pélissier.

Avant-propos : Philippe Bonfils et Jacques Walter.

Mise en forme éditoriale : Mélanie Bourdaa et Norbert Congard.

ISBN : 978-2-914872-00-3

9 782914 872003

Illustration de couverture : pixabay.com. Réalisation de couverture : www.atelier-congard.fr