

Humanités numériques et archives : la longue émergence d'un nouveau paradigme

Digital Humanities and Archives: The Lengthy Emergence of a New Paradigm

Frédéric Giuliano

Volume 65, Number 2, April–June 2019

Techno, techno, techno...

URI: <https://id.erudit.org/iderudit/1063788ar>

DOI: <https://doi.org/10.7202/1063788ar>

[See table of contents](#)

Publisher(s)

Association pour l'avancement des sciences et des techniques de la documentation (ASTED)

ISSN

0315-2340 (print)

2291-8949 (digital)

[Explore this journal](#)

Cite this article

Giuliano, F. (2019). Humanités numériques et archives : la longue émergence d'un nouveau paradigme. *Documentation et bibliothèques*, 65(2), 37–46.
<https://doi.org/10.7202/1063788ar>

Article abstract

The purpose of this article is to describe and clarify research as it applies to the digital humanities and how their contribution and potential can have an impact on archival and library practices. The author outlines the conceptual and technological attributes of the digital humanities and the challenges faced by information professionals associated with these projects as well as the emerging role they are called upon to play.

HUMANITÉS NUMÉRIQUES ET ARCHIVES: LA LONGUE ÉMERGENCE D'UN NOUVEAU PARADIGME

Frédéric Giuliano

Archiviste à l'Université McGill

frederic.giuliano@mcgill.ca

RÉSUMÉ | ABSTRACT

Cet article vise à démystifier le domaine de recherche que sont les humanités numériques en révélant sa contribution et l'immense potentiel de ses approches pour la pratique archivistique et bibliothéconomique. L'auteur y présente une réflexion sur les apports conceptuels et technologiques des humanités numériques tout en soulignant les défis que doivent relever les professionnels de l'information associés à ces projets, mais aussi les nouveaux rôles que ces derniers sont appelés à y jouer.

Digital Humanities and Archives: The Lengthy Emergence of a New Paradigm

The purpose of this article is to describe and clarify research as it applies to the digital humanities and how their contribution and potential can have an impact on archival and library practices. The author outlines the conceptual and technological attributes of the digital humanities and the challenges faced by information professionals associated with these projects as well as the emerging role they are called upon to play.

Introduction

« *Le médium, c'est le message.* »

(McLuhan, 1968)

Alors qu'il est maintenant scientifiquement démontré que les langues que nous parlons modifient notre façon de penser, de percevoir le monde et nos capacités cognitives, n'en va-t-il pas de même des technologies que nous utilisons? C'est à tout le moins la thèse du théoricien des médias, Marshall McLuhan (1968), qui affirme que la nature d'un médium est plus importante que le sens ou le contenu du message, mais aussi que le vrai message, tant en réalité qu'en pratique, c'est le médium lui-même, « c'est-à-dire, tout simplement, que les effets d'un médium sur l'individu ou sur la société dépendent du changement d'échelle que produit chaque nouvelle technologie, chaque prolongement de nous-mêmes, dans notre vie ».

Suivant cette logique, s'il est vrai que les inventions et innovations technologiques que sont l'ordinateur personnel et Internet ont profondément bouleversé nos disciplines, il en est tout autant de cet étrange domaine de recherche que sont les humanités numériques qui imposent doucement, depuis déjà 70 ans, leurs pratiques et outils dans les milieux de recherche. Telle une petite révolution, ces pratiques et

outils qui utilisent la puissance de l'ordinateur pour traiter et analyser des quantités de plus en plus grandes de données amènent de nouvelles façons de chercher, de gérer et de traiter l'information. En conséquence, les archivistes et bibliothécaires impliqués avec ces chercheurs sont forcés de penser et percevoir différemment leurs collections afin de répondre à ces nouveaux besoins. Ainsi, les dépôts de connaissances que sont les centres d'archives et les bibliothèques se transforment graduellement : lieux de conservation et de diffusion de connaissances et de savoirs, ils deviennent aussi des lieux de conservation et de diffusion de données interprétables par l'ordinateur.

Cet article présente l'évolution des humanités numériques en soulignant plus directement ses apports conceptuels et technologiques. Il s'en suit une réflexion sur les impacts profonds que ces technologies et approches ont eu sur la pratique archivistique et bibliothéconomique, qui présente les défis posés par les humanités numériques pour les professionnels de l'information impliqués dans de tels projets. Ce texte a ainsi pour but de contribuer à démystifier le domaine de recherche que sont les humanités numériques en révélant sa contribution et l'immense potentiel de ses approches pour la pratique archivistique et bibliothéconomique.

Les humanités numériques: la petite histoire d'une longue évolution

« Les humanités numériques ne sont pas seulement à propos du numérique ou seulement à propos des humanités, mais se définissent plutôt par les opportunités et les défis qui se posent par la conjonction des termes humanités et numériques qui forment ainsi une nouvelle réalité. »

(Burdick, Drucker, Lunenfeld, Presner et Schnapp, 2012)

Les débuts

À la fin des années 1940, le père jésuite Roberto Busa (1980) eut l'idée d'utiliser l'ordinateur dans le cadre de ses recherches sur Thomas d'Aquin. Avec l'assistance de Thomas J. Watson d'IBM, Busa et son équipe créèrent un index d'environ 11 millions de mots en latin médiéval utilisés par Thomas d'Aquin et d'autres auteurs lui étant associés. Une tâche monumentale s'il en est! La totalité du texte fut graduellement transférée sur des cartes perforées dans l'optique de générer automatiquement un index de chacun des mots du corpus. Cet *Index Thomisticus* fut publié en 1974, en 56 volumes. Les travaux de Busa sont considérés comme étant les premiers à avoir utilisé l'informatique pour faire des recherches en sciences humaines. À juste titre, Busa voyait en l'informatique un outil facilitant les pratiques de recherche en sciences humaines et sociales, grâce à la puissance computationnelle (c'est-à-dire de calcul) de l'ordinateur qui accélère la fastidieuse tâche d'analyse manuelle des textes, tout en permettant « l'amélioration de la qualité, de la profondeur et de l'extension de la recherche et pas simplement la réduction du temps et de l'effort humain » (Busa, 1980). Cette utilisation de l'informatique ne vient cependant pas changer les pratiques de recherche, mais elle les simplifie, les automatise et les rend plus rapides (Sinatra et Vitali-Rosati, 2014). C'est en ce sens qu'on parle alors d'humanités computationnelles (*humanities computing*), ou pour reprendre les termes exacts de l'époque: *literary and linguistic humanities computing*. Évidemment, ces outils informatiques étaient alors peu répandus, très coûteux et utilisés seulement par une petite communauté de chercheurs. La démocratisation de ces pratiques ne viendra que quelques décennies plus tard avec l'arrivée de l'ordinateur personnel, puis d'Internet.

Les années 1960-1970

Les décennies qui suivent voient de nombreux travaux similaires à ceux du père Busa faire leur apparition en Europe et en Amérique du Nord, comme le projet *Brown Corpus for Use on Digital Computers*, qui vise la constitution d'un corpus de la langue anglaise courante, et dont la première version compte environ un million de mots, ou encore le projet *British National Corpus* constitué d'environ 100 millions de

mots, un volume impensable sans la technologie informatique (Burnard, 2012).

Cependant, un des projets les plus notables de cette période fut certainement celui de Theodore F. Brunner de l'Université de Californie. L'idée était de constituer une banque de données de tous les textes anciens, de Homère jusqu'au VI^e siècle, ce qui représente environ 70 millions de mots. Financé en grande partie par David Hewlett (fils du fondateur de Hewlett Packard) qui rencontra Brunner lors d'un des cours du soir sur la littérature grecque dispensé par ce dernier, et à l'occasion duquel il évoquait les possibilités de recherche sur les fonds littéraires grecs, le *Thesaurus Linguae Graecae* (TLG) allait avoir un impact majeur sur la communauté scientifique (Burnard, 2012).

Le développement progressif de tels outils électroniques a incité de nombreux chercheurs à en étudier leurs impacts: Peter Lyman (1984) et Karen Ruhleder (1995) examinent tous les deux comment la conversion du papier aux sources électroniques a affecté la prise en charge des données, des processus de travail et la réflexion qui en découle. Alors que Lyman note que la création d'outils électroniques a changé les relations sociales et le sens associé à la notion de travail, Ruhleder étudie comment le TLG influence le processus de travail des universitaires œuvrant dans le domaine des études classiques. Au moment de son étude, en 1993, le TLG jouissait d'une réputation enviable au sein du milieu des études classiques: près de 1490 cédéroms contenant le TLG étaient en circulation et huit distributeurs offraient l'outil sur différentes plateformes d'exploitation. En se basant sur 60 entrevues non structurées et l'observation participative d'utilisateurs du TLG au sein de départements d'études classiques, Ruhleder souhaitait mesurer l'impact de la technologie sur leur travail. L'étude incluait les logiciels de traitement de texte, du courriel, l'utilisation d'une banque de données en ligne, et bien sûr le TLG. Elle a ainsi démontré que la facilité de recherche qu'amenait l'utilisation de ces outils avait profondément changé la façon dont les universitaires approchaient leur corpus, plus particulièrement ceux qui se trouvaient en début de carrière. Avec l'apparition des algorithmes de recherche apparut une nouvelle pratique d'érudition permettant par exemple à des étudiants diplômés de poser des questions pour lesquelles seule une lecture extensive du corpus permettait de trouver une réponse (Dalbello, 2011). Ruhleder (1995) conclut son étude en insistant sur l'effet profond et restructurant de l'utilisation du TLG sur la production de savoir dans le domaine des études classiques. Ainsi, un chercheur est maintenant en mesure de réaliser un travail intellectuel plus intéressant dans la mesure où il ne perd plus son temps à l'apprentissage extensif du corpus. Cette nouvelle forme d'accessibilité aux sources premières a de ce fait transformé la façon dont une recherche peut être conduite. On assiste dès lors à un changement en profondeur de la pratique

érudite, à un changement de paradigme, et à un passage de la maîtrise érudite d'un corpus vers la construction de problèmes (constructivisme heuristique). L'évolution technologique des années à venir devrait accélérer encore cette tendance, tout en continuant d'accroître l'accessibilité aux corpus.

La popularité de ces travaux et l'intérêt grandissant pour cette approche interdisciplinaire conduisent à la fondation en Angleterre de l'*Association for Literary and Linguistic Computing* en 1972 et à la création de la revue *Literary and Linguistic Computing* qui lui est associée en 1976, revue qui existe encore aujourd'hui. En outre, toujours durant cette période, deux associations professionnelles font leur apparition : l'*Association for Literary and Linguistic Computing* (ALLC) en 1973, et l'*Association for Computers and the Humanities* (ACH) en 1978. Durant les décennies 1960 et 1970, on assiste par ailleurs à une multiplication des centres et des programmes de formation universitaire. Une institutionnalisation de la discipline attestée par le passage de la *literary and linguistic computing* aux *humanities computing* (Burnard, 2012) confirme que les pratiques de recherche assistée par ordinateur dans le domaine des sciences humaines et sociales ne sont plus l'apanage de quelques chercheurs isolés, mais bien une véritable approche interdisciplinaire, partagée par des communautés de recherche structurées et organisées (Sinatra et Vitali-Rosati, 2014).

En somme, ce que cette première période des humanités computationnelles a révélé est, premièrement, l'identification d'un problème (encore présent de nos jours), à savoir le besoin de voir au-delà de la chaîne de caractères des mots pour s'attarder, d'une manière efficace, aux différences de graphie de ces derniers, aux multiples manuscrits et à la lemmatisation (Hockey, 2004). Deuxièmement, au-delà de l'amélioration des systèmes et de la multiplication des organes de diffusion que sont les conférences thématiques et revues savantes, c'est le besoin de logiciels, de normes pour l'archivage et de conservation des textes qui prévaut (Hockey, 2004). C'est exactement ce que les deux décennies suivantes amèneront aux humanités numériques.

Les années 1980-1990

L'adoption de la nouvelle génération d'outils électroniques que sont l'ordinateur personnel, le courriel et Internet (à partir du milieu des années 1990) a eu des effets majeurs sur les sciences humaines et sociales. De ces trois innovations, la plus marquante est probablement celle du courriel, car sa création permet le développement des listes de diffusion, ce qui facilite grandement l'échange et le partage d'idées, au cœur même de la pratique scientifique. En ce sens, le courriel constitue le changement de médium le plus significatif. La création du courriel permet le développement des listes de diffusion, ce qui facilite grandement

l'échange et le partage d'idées, au cœur même de la pratique scientifique. C'est pour répondre à ce besoin de rester en contact sur une base régulière que Willard McCarty, professeur en humanités computationnelles a fondé, en 1987, la liste de diffusion *Humanist* (Hockey, 2004 ; Burnard, 2012), dont l'objectif « est de fournir un forum de discussion sur des questions intellectuelles, académiques, pédagogiques et sociales et un échange d'informations entre les participants¹ ». Cette liste de diffusion est d'ailleurs toujours bien vivante et encore éditée par son fondateur, Willard McCarty.

Une autre avancée majeure des humanités computationnelles est assurément à chercher du côté des initiatives de normalisation de l'encodage des données. La problématique était de « rendre possible la mutualisation des données en apportant une solution à l'immense variété des manières d'encoder » (Burnard, 2012). La *Text Encoding Initiative* (TEI) est venue répondre à ce besoin, tel un modèle unique et encyclopédique de représentation des éléments signifiants d'un texte qui a rapidement fait consensus au sein de la communauté des chercheurs. La TEI sert également de modèle aux développeurs de la norme *Encoded Archival Description* (EAD) relative à la création des instruments de recherche du milieu archivistique et des collections spéciales (Hockey, 2004).

C'est ce même problème ou nécessité de balisage des éléments d'un texte dans l'univers numérique qui est à la base de la création du code HTML, perçu par certains comme une version édulcorée de la TEI, dont l'impact sur l'érudition numérique a été énorme. Aujourd'hui, la TEI est reconnue à l'échelle internationale comme un outil extrêmement important, à la fois pour la conservation à long terme des données électroniques et comme moyen de soutenir l'utilisation efficace de ces données dans de nombreux domaines. C'est le schéma d'encodage de choix pour la production d'éditions critiques et savantes de textes littéraires, d'ouvrages de référence savants et de grands corpus linguistiques, ainsi que pour la gestion et la production de métadonnées détaillées associées à des collections de textes électroniques et de patrimoine culturel de nombreux pays.

Un autre projet fondateur dans l'histoire des humanités numériques est certainement le projet Perseus de bibliothèque numérique qui a débuté en 1985 et rassemblait des ressources relatives aux études classiques. Le projet Perseus avait comme prémisse – quelque peu polémique – que les bibliothèques numériques mettraient à jour de nouvelles méthodes grâce auxquelles on pourrait poser de nouvelles questions à propos de nouvelles idées qu'il aurait été impossible d'explorer sans elles (Crane et Wulfman, 2003 ; Dalbello, 2011). En effet, c'est ce que de nombreuses

1. dhhumanist.org

études sur son impact ont démontré (Marchionini, 2000). La promotion de l'utilisation des sources primaires par opposition aux sources secondaires pour les étudiants universitaires a amené une approche heuristique et didactique constructiviste.

Au final, le projet Perseus a fourni un modèle pour les humanités numériques en termes de contenu et d'utilisation. Plus encore, il a démontré que la technologie ne change pas seulement la manière dont les disciplines fonctionnent au sens général, mais qu'elle influence aussi directement l'infrastructure conceptuelle d'un champ de recherche. Cette idée est conforme à la principale conclusion de l'étude du PDL, à savoir que l'expérience de recherche change qualitativement grâce à la recherche et à l'accessibilité de sources primaires et d'environnements multimédias, car ces derniers permettent un engagement constructiviste avec des connaissances et des textes existants, et que les humanistes perçoivent leur propre travail de manière analogue à un laboratoire pour la construction heuristique (Dalbello, 2011).

Encore une fois, en écho à la théorie de McLuhan, l'apport de la technologie change la relation entre l'utilisateur et le contenu. De fait, elle suscite un nouvel engagement envers les sources premières – dont font partie les archives – qui se traduit notamment par une lecture plus active du contenu. La technologie des liens hypertextuels et celle des moteurs de recherche, en particulier, a orienté ce nouvel engagement ou nouvelle relation aux archives en permettant différents types de recherche contextuelle. Celles-ci ont révélé à leur tour de nouvelles interactions jusque-là invisibles dans les interactions traditionnelles (Latham, 2004). Dans une interprétation sobre de la science cognitive, on peut alors parler de cet état comme étant une surcharge cognitive engendrant la passion plutôt que de lasser ou de frustrer le lecteur (Dalbello, 2011).

Des années 2000 à aujourd'hui

Le passage des *humanities computing* vers les *digital humanities* se concrétise en 2004 avec la publication de l'ouvrage *Companion to Digital Humanities*, qui démontre la vitalité et le caractère interdisciplinaire des humanités numériques (Sinatra et Vitali-Rosati, 2014). C'est en fait le début de la constatation de la révolution que ces outils ont engendrée en changeant profondément le concept même des disciplines humanistes. La naissance d'Internet et sa diffusion rapide au milieu des années 1990 ont accru l'impact de cette révolution. Le changement de nom marque ainsi une étape importante, car il confirme la transformation d'une approche méthodologique en une discipline à proprement parler (Sinatra et Vitali-Rosati, 2014). L'arrivée d'Internet est donc en ce sens un moment charnière dans le développement des humanités numériques, car plus qu'un simple

outil ou une simple technique modifiant nos pratiques de recherche, il devient un objet de recherche et modifie notre façon de voir le monde.

Vers une définition holistique des humanités numériques

Définir les humanités numériques de façon stable et robuste pose encore un défi de taille. En effet, si l'évolution longue et variée que nous avons détaillée plus avant explique en partie cette difficulté, le concept même varie encore et toujours selon à qui l'on s'adresse. Cette difficulté de définition des humanités numériques reflète indéniablement la vigueur des débats entourant la discipline. Pour les uns, c'est un signe de vigueur ; pour les autres, une fragilité telle qu'elle remet en question jusqu'à l'existence même des humanités numériques comme domaine d'études. Le dossier piloté par Fabien Granjon (2016) sur la critique des humanités numériques de la revue *Variations* est des plus intéressants à cet égard.

Au-delà de ce débat, ce que l'évolution des humanités numériques nous démontre aujourd'hui, c'est que celles-ci ne sauraient être définies uniquement comme étant la simple utilisation d'outils informatiques pour la recherche ou la communication en sciences humaines et sociales (Burdick, Drucker, Lunenfeld, Presner et Schnapp, 2012). En effet, en plus d'être une méthode d'analyse informatique appliquée aux sciences humaines et sociales, les humanités numériques transcendent cet aspect technique et peuvent être pensées davantage comme un regard global posé sur les changements culturels déterminés par le numérique. De là, il n'y a qu'un pas à franchir vers une conversion numérique ouvrant la porte à un concept d'humanisme numérique tel que défini par Milad Doueïhi (Sinatra et Vitali-Rosati, 2014).

Aujourd'hui, les humanités numériques font référence à un nouveau mode d'érudition et d'organisation institutionnelle collaborative, transdisciplinaire et computationnel, engagé dans la recherche, l'enseignement et la publication (Burdick, Drucker, Lunenfeld, Presner et Schnapp, 2012). Les humanités numériques ne sont donc pas seulement à propos du numérique ou seulement à propos des humanités, mais se définissent plutôt par les opportunités et les défis qui se posent par la conjonction des termes humanités et numériques qui forment ainsi une nouvelle réalité. Comme le disent Sinatra et Vitali-Rosati (2014), « dans le débat actuel, on essaie de ne pas penser les humanités numériques comme une discipline, mais plutôt les envisager comme une approche globale, transdisciplinaire, adoptant une attitude et un point de vue sur la recherche qui devraient impliquer l'ensemble des chercheurs en sciences humaines et sociales ».

Un nouveau paradigme

« Nous sommes face à un changement de paradigme dans la façon d'agencer les contenus et, par ce fait même, à une nouvelle conception du savoir et de sa circulation dans la société. »

(Sinatra et Vitali-Rosati, 2014)

Le texte de Vannevar Bush, *As We May Think*, écrit en 1945, et celui de Ted Nelson, *A File Structure for the Complex, the Changing and the Intermediate*, écrit en 1965, annoncent bien davantage qu'un simple changement technologique. C'est en effet un véritable changement de paradigme de la structuration du savoir qui se développe depuis la fin des années 1940. Les humanités numériques sont une discipline chargée de réfléchir aux outils technologiques sur lesquels les sciences humaines et sociales devraient s'appuyer, mais aussi de produire l'appareil théorique pour interpréter les structures conceptuelles fondamentales qui caractérisent notre culture actuelle (Sinatra et Vitali-Rosati, 2014).

Pour les archivistes et bibliothécaires, cela se traduit par le fait que, plus que jamais, les archives et collections spéciales sont un laboratoire pour les humanités numériques. Ces professionnels sont à la fois collaborateurs, chercheurs, éditeurs, rédacteurs et organisateurs. Ce n'est pas étonnant lorsqu'on pense à la richesse que contiennent les différentes collections d'archives, richesse malheureusement presque toujours inexploitable par les moteurs de recherche. Comme nous l'avons mentionné plus haut, l'édition savante et l'organisation numérique des archives ont été au centre du développement des humanités numériques. De nouveaux modèles de recherche et d'érudition en ont émergé. Quelques projets de collection numérique d'archives ont fait école, et ce, dès les années 80, notamment avec les projets *The Rossetti Archive* et *The William Blake Archive* réalisés par l'unité de recherche *The Institute for Advanced Technology in the Humanities* (IATH) de l'Université de Virginia. Au cours des années 90, d'autres projets similaires ont suivi, tels que *Dickinson Electronic Archives* de l'Université du Maryland et *The Walt Whitman Archive* du *Center for Digital research in the humanities* de l'Université du Nebraska. Ces collections savantes ont été élaborées selon un processus éditorial minutieux. Elles sont représentatives des types de recherche pratiqués dans les environnements numériques lorsque la création d'artefacts technologiques, tels que des bases de données, accompagne la construction de significations relatives à la reproduction de textes par le biais de technologies et de représentations numériques.

L'optimisme avec lequel les technologues humanistes voient la création d'archives hypertextuelles et reconnaissent l'effet de démocratisation de la technologie en permettant de donner accès à des matières premières, ne se limite pas à une posture rhétorique. Il ne s'agit pas en effet d'une idéalisation

d'un nouvel âge de lecture pour la masse des lecteurs en ligne ordinaires. Néanmoins, l'optimisme et l'enthousiasme suscités par les nouvelles possibilités de repenser les textes et la création sont en avance sur le travail de développement réel et actuel des humanités numériques. En dépit de l'omniprésence de la technologie numérique, ses effets sur les sciences humaines sont également limités, car de profondes transformations des infrastructures et des programmes sont nécessaires pour s'adapter aux nouvelles formes de création de connaissances et d'apprentissage (Dalbello, 2011).

De l'information aux données

Dans son aspect plus technique ou pratique, il est possible d'englober la pratique des humanités numériques en dix façons de travailler. Ce sont la numérisation, la participation collaborative (*crowdsourcing*), les bases de données, la conservation numérique (*digital curation*), les textes, l'édition, la visualisation, la géolocalisation, les jeux et le code (Poole, 2017). Bien entendu, il existe une pléthore d'outils sous-jacents à ces pratiques ou façons de travailler.

Les archivistes et bibliothécaires auraient beaucoup à gagner à inclure ces technologies et moyens de recherche des humanités numériques dans leur pratique professionnelle, que ce soit en matière d'acquisition, de traitement et surtout de diffusion de leurs collections. Soulignant à juste titre la difficulté qu'ont les utilisateurs d'archives à comprendre et utiliser les collections d'archives et encore bien davantage le jargon archivistique ou bibliothéconomique, Anne Bahde (2017) a tenté l'expérience de marier la visualisation de données aux instruments de recherche en ligne. Elle a ensuite soumis ses modèles à de petits groupes d'utilisateurs (*focus groups*). La réaction de ces derniers quant aux outils de visualisation confirme les théories avancées par Ruhleder et Lyman sur l'engagement des utilisateurs dans la recherche. Plusieurs utilisateurs ont indiqué que le format de la chronologie améliorerait leur capacité à se souvenir de l'ordre des événements et ont indiqué que cette fonctionnalité serait particulièrement utile pour les grandes collections ou pour les personnes et les organisations dont l'historique est complexe (Bahde, 2017). Des experts, notamment des professeurs d'histoire ou ayant intégré des recherches historiques dans leurs cours, ont déclaré qu'ils souhaitaient que leurs étudiants voient cet outil. Ils ont évoqué son potentiel de curiosité, mais peu ont déclaré qu'ils utiliseraient la chronologie et la carte dans leurs propres recherches. Ce qui est néanmoins ressorti collectivement de leurs impressions et de leurs idées d'amélioration est la vision d'un environnement de recherche archivistique hautement flexible et personnalisable qui s'adapte aux besoins individuels des utilisateurs en matière de recherche, aux comportements de recherche d'information et au niveau de compétence, et qui permet divers niveaux d'interaction et de contrôle d'information (Bahde, 2017).

En représentant les données archivistiques de différentes manières, que ce soit dans les chronologies, les graphiques réseau, la modélisation de sujets ou les cartes à cases (*tree-maps*), nous pouvons inviter à de nouvelles interprétations de ces collections, encourager l'expérimentation et le développement de nouvelles méthodes de découverte archivistique en plus d'implanter de nouvelles sources d'inspiration pour nos utilisateurs. Le jumelage de nouveaux paradigmes de visualisation avec les concepts de classification et de description récemment proposés peut offrir des possibilités intéressantes d'accès aux documents d'archives dans les décennies à venir.

À l'instar du projet de Bahde, de nombreux autres projets ont démontré les effets positifs de l'utilisation de modèles de visualisation, des plus simples comme les diagrammes ou graphiques statiques pour la comparaison de données aux plus dynamiques comme des générateurs de connaissances visuels (*knowledge generating visualizations*). En effet, la visualisation de données permet de rendre certains aspects d'un phénomène plus visibles que d'autres (Berry, 2012). Ces visualisations permettent aux utilisateurs de participer au processus d'analyse, de sonder et d'explorer plus avant, de remixer et de jouer avec les données pour parvenir à une nouvelle compréhension (Bahde, 2017). Les données archivistiques invitent ainsi à des possibilités infinies d'analyse visuelle. La visualisation des données d'archives peut permettre au chercheur et à l'archiviste de comparer, mettre en correspondance, reconnaître, distinguer, organiser, combiner, construire et organiser des données sur une constellation de points de données, d'une manière que les instruments de recherche textuels traditionnels ne peuvent faire (Bahde, 2017). C'est exactement ce à quoi référaient les études de Karen Ruhleder sur les notions de *close reading*

et de *distant reading*. Ainsi, présenter des données d'une façon qui permette d'accomplir une lecture distancée peut inciter les utilisateurs à identifier des modèles temporels ou spatiaux, des relations entre des collections ou des proportions entre des catégories de données. De telles informations peuvent habiliter le chercheur à effectuer différentes étapes du processus de recherche archivistique, qu'il s'agisse de choisir un sujet de recherche, d'identifier des collections pertinentes ou d'établir un contexte.

Bien qu'un article de Robert B. Allen (2005) suggérait diverses approches possibles de techniques de visualisation appliquées au domaine archivistique, incluant des lignes de temps, de la cartographie, ainsi que des réseaux et des outils pour explorer et articuler les hiérarchies archivistiques institutionnelles, plus particulièrement gouvernementales, Bahde (2017) souligne que bien peu d'outils et de recherches expérimentant ces outils ont pour l'instant émergé.

Néanmoins, un projet de visualisation de données a retenu plus particulièrement notre attention. Il s'agit de *The Visible Archive Project*, réalisé par le D^r Mitchell Whitelaw et les *National Archives of Australia*, lancé en 2009, et qui propose une visualisation de leur collection d'archives. En appliquant des techniques de visualisation de pointe à de vastes ensembles de données archivistiques, le D^r Whitelaw a mis au point de nouvelles méthodes d'exploration et de compréhension des relations qui donnent aux enregistrements individuels leur contexte. Ses recherches ont suscité un intérêt international considérable et suggèrent de nouvelles possibilités intéressantes d'accès du public aux collections du patrimoine culturel via le *Series Browsers*². Celui-ci offre une visualisation à l'aide d'une carte à cases qui présente les relations de près de 65 000 séries archivistiques en utilisant des variations de couleurs et de formes.

FIGURE 1

Exemple d'une carte à cases générée à l'aide du *Series Browsers*.



2. visiblearchive.blogspot.com/2009/07/bi-directional-series-links.html

Le projet offre également la possibilité de réaliser une recherche interactive et des nuages de mots démontrant les connexions entre certains mots clés à travers la collection grâce au *A1 Explorer*³. Plus bas, on peut voir un exemple de la manière dont le nuage de mots, les cooccurrences et l'histogramme peuvent être combinés pour révéler des modèles dans le jeu de données et provoquer des découvertes dans le contenu de la série. En se concentrant sur « Darwin », on remarque une forte hausse dans l'histogramme en 1937. Dans ce cas, les cooccurrences et la distribution des termes donnent un aperçu précis de ce que sont les éléments de cette pointe : un cyclone. En cliquant sur Darwin, le nuage de mots révèle même le mois de l'événement et, à nouveau, la liste des éléments montre une confirmation détaillée du motif.

cette nouvelle façon d'être et de penser, suivant l'évolution des nouvelles technologies, amène sa part de défis pour les professionnels de l'information et leurs institutions. La jeune histoire des humanités numériques met en lumière le rôle fondamental des archives dans la naissance et l'évolution de la discipline. Autant les archivistes et bibliothécaires ont embrassé le Web 2.0 et Internet pour la diffusion de leurs collections, autant ils sont restés majoritairement en marge de tout ce qui a trait à l'utilisation de la puissance de l'ordinateur en matière de gestion et de traitement de leurs collections. Comment l'expliquer ?

Il y a tout d'abord une problématique importante qui entoure le développement des outils spécialisés supportant les humanités numériques. Ces outils sont essentiels à l'existence même de la discipline ; sans ordinateur, pas de

FIGURE 2

Exemple d'un nuage de mots jumelé à un historiogramme généré par le A1 Explorer.



L'utilisation des humanités numériques rapproche donc les archives des besoins des utilisateurs, notamment par la normalisation du langage, celui des données structurées et non d'une logique archivistique comprise seulement par quelques initiés et les archivistes eux-mêmes.

Les défis des humanités numériques

Tel qu'énoncé en prémisses à ce texte, les développements technologiques influent notre façon d'être et de penser, comme l'affirmait Marshall McLuhan (1968). Plus encore,

numérique, et sans outils, pas d'humanités numériques. Mais aussi vitaux que soient le développement et l'existence de ces outils, six problèmes demeurent. Premièrement, très peu de développeurs d'outils prennent le temps de mener une étude sur les possibilités et les risques inhérents au projet (*environmental scans*) avant de débiter. Deuxièmement, une absence de gestion ou une mauvaise gestion de projet et de ressources humaines peuvent gâcher le développement d'outils. Troisièmement, trop d'outils, bien qu'opérationnels, se retrouvent orphelins, car ils ne sont pas rattachés à un centre de recherche ou à une institution. Quatrièmement, les développeurs abandonnent, remplacent ou délaissent leurs outils une fois la recherche terminée.

3. visiblearchive.blogspot.com/2009/08/exploring-a1-items-to-documents.html

Cinquièmement, beaucoup d'outils sont développés par et pour des experts, et manquent de la convivialité nécessaire à un utilisateur non initié, ce qui ne favorise pas l'adhésion par un groupe plus large d'utilisateurs. Sixièmement, les développeurs n'investissent pas suffisamment dans le maintien de leur outil, si bien qu'ils n'en sont souvent que les seuls utilisateurs (Poole, 2017).

Outre la problématique entourant le développement des outils des humanités numériques, d'autres défis se posent : l'interdisciplinarité, la collaboration et la gestion de projet (Poole, 2017). En effet, la pratique de l'interdisciplinarité est difficile, d'abord parce que les collaborateurs ont une compréhension différente d'un problème de recherche donné, et des approches différentes quant à la façon de conduire la recherche. Ensuite, parce que les parties prenantes au projet ont tendance à développer des objectifs durant la conduite du projet, et non en amont. Ce caractère évolutif intrinsèque aux humanités numériques cause de sérieux problèmes (Thomas III, 2016).

L'un des aspects fondamentaux de la pratique des humanités numériques est assurément son aspect collaboratif. Cette collaboration est centrée sur une compréhension commune du langage et de la terminologie, des méthodes, des théories, des tâches et des flux de travail, des valeurs, des objectifs et des résultats. Pour réaliser une collaboration réussie, les équipes de recherche doivent cultiver la confiance entre leurs membres et arriver à un consensus par une communication respectueuse et efficace. Plus concrètement, ils doivent planifier des réunions, répartir des tâches, mais aussi partager des informations et des documents. Indépendamment de la discipline, les défis incluent les personnalités, la communication et la coordination, l'égalité des contributions, la responsabilisation et la responsabilité, la formation et la répartition d'un financement cohérent (Poole, 2017).

Le caractère interdisciplinaire et collaboratif des humanités numériques implique une approche par gestion de projets de toute entreprise. La gestion des ressources humaines, la planification, la création de contenu, le développement et la maintenance technique, la préservation, la diffusion et le stockage sont autant de besoins de gestion qu'il importe de gérer adéquatement. Le travail en projet collaboratif reste variable en ce qui concerne le financement, les délais, la dynamique d'équipe et les problèmes abordés, et pour compliquer les choses, la plupart des professionnels en humanité numérique manquent de préparation formelle à la gestion de projet. Les échecs les plus fréquents incluent notamment le manque de planification, l'absence de produits livrables et la perte de leadership (Poole, 2017). Les compétences recherchées en gestion de projet comprennent l'organisation, la planification et le suivi, la définition des priorités, l'administration des subventions, la

gestion des ressources humaines et la résolution des conflits. Enfin, à mesure que le projet avance, les parties prenantes doivent prendre en compte des problèmes tels que le glissement de la portée du projet et des fonctionnalités de l'outil (par exemple, par l'ajout de fonctionnalités gratuites) et fournir d'autres informations et formations aux utilisateurs (Reed, 2014). «Un projet sans responsabilité, sans connexion, sans critique remplit simplement un autre caveau dans le cimetière de projets d'humanités numériques», prévient Fred Gibbs (2011).

Le rôle des bibliothécaires et archivistes en humanités numériques

Cela étant dit, ces défis de la pratique de l'interdisciplinarité, de la collaboration et de la gestion de projet ne sont pas des choses nouvelles pour les bibliothécaires et archivistes : on pourrait pratiquement dire qu'ils font partie intégrante de leur pratique professionnelle et de leur discipline. De ce fait, ne sont-ils pas les mieux placés pour y faire face ? Comme le souligne à juste titre Klein (2015) : «Les humanités numériques sont un domaine en pleine croissance aux carrefours de l'informatique et des disciplines des sciences humaines et des arts, des domaines interdisciplinaires de la culture et de la communication, ainsi que des professions de l'éducation et des sciences de la bibliothèque et de l'information.»

Lorsqu'on s'attarde plus longuement aux disciplines des humanités numériques et des sciences de l'information, il est frappant de constater à quel point celles-ci sont complémentaires. Tout d'abord, la matière première des deux disciplines est l'information et les documents. Ensuite, les deux concentrent leur pratique et leurs recherches sur des sujets similaires : la gestion et le développement des collections, la recherche et l'accès à l'information, les métadonnées, la description, la classification, le développement d'ontologies et de taxonomies, l'accès libre, le Web des données (*linked data*), la préservation, l'expérience utilisateur, le développement d'interfaces, le *big data*, le forage de texte, etc. Les deux domaines incluent dans leurs programmes de formation des sujets de l'autre domaine, tout comme dans les conférences et revues. Tous les deux embrassent l'interdisciplinarité et les technologies et techniques permettant l'exploration scientifique des données. En résumé, les deux domaines reposent sur les mêmes bases théoriques et conceptuelles qui sous-tendent leurs activités pratiques (Robinson, Priego et Bawden, 2015).

Plus précisément, les rôles que peuvent occuper archivistes et bibliothécaires au sein des humanités numériques sont multiples, comme le souligne une littérature somme toute assez vaste. À titre illustratif, dans une étude sur le rôle joué par les archivistes et bibliothécaires au sein de 11 projets subventionnés en humanités numériques, Poole et

Garwood (2018) ont pu leur établir huit rôles principaux : l'aide pédagogique aux utilisateurs, l'offre d'espace physique et virtuel, la découverte et l'accès aux ressources électroniques, la gestion et la création des métadonnées, la gestion de projet, le maintien à long terme des projets, la préservation numérique et l'hébergement, et le stockage à long terme des données. Ces huit domaines d'intervention indiquent que les bibliothécaires et les archivistes ont joué un rôle crucial dans la réussite des projets. Même s'ils n'étaient pas des membres officiels des équipes de projet, ils assumaient un soutien crucial dans presque tous les cas (Poole et Garwood, 2018). Le fait que les compétences des archivistes et bibliothécaires demeurent somme toute mal connues chez les chercheurs engagés dans des projets d'humanités numériques nuit assurément à l'atteinte du plein potentiel de la collaboration entre les deux domaines. Une offre plus importante de formation en humanités numériques dans les programmes de formation en sciences de l'information, et inversement, aiderait à l'atteinte de ce plein potentiel collaboratif.

Conclusion

Tout doucement, depuis la fin des années 1940, les humanités numériques, au fil des évolutions technologiques, ont profondément bouleversé l'agencement, la création et notre conception du savoir, mais aussi notre façon d'interagir avec celui-ci. Ces changements profonds ont permis l'émergence d'un nouveau paradigme dans le domaine des sciences de l'information.

À l'instar des humanités numériques qui doivent développer une réflexion sur la façon dont les outils numériques changent la recherche en sciences humaines et sociales, les archivistes et bibliothécaires doivent investir pleinement le champ des humanités numériques et réfléchir aux façons d'utiliser les outils numériques dans leur discipline, ce qui nécessite de s'approprier ce nouveau paradigme et de réfléchir différemment à leurs disciplines. L'importance des archives et, pour les archivistes et les bibliothécaires, de prendre corps avec ces approches pour repenser leur discipline, est fondamental pour leur évolution. Nombreux sont ceux qui dans la profession maintiennent l'espoir que nous pouvons créer une ère d'accès complet aux documents d'archives en offrant aux utilisateurs une exploration de recherche, des découvertes et même des environnements de jeu qui rendent les documents historiques facilement accessibles à de nouveaux publics. Nous pouvons bouleverser le modèle traditionnel de la force et de l'inertie de l'instrument de recherche en archivistique et permettre aux utilisateurs de participer à la création de connaissances, de manipuler les systèmes en fonction de leurs propres besoins et de créer de nouveaux liens entre les matériaux. Bien que cela nécessite un changement dans la manière dont nous examinons les données archivistiques, une volonté de redéfinir nos compétences de base, une volonté d'établir une redistribution importante de la main-d'œuvre et des ressources vers la programmation, la conception d'interfaces et les études de convivialité, les récompenses offertes à nos utilisateurs et à l'érudition seraient inestimables. Et de là, nous pourrions vraiment exploiter pleinement le potentiel des humanités numériques au bénéfice de nos disciplines.

SOURCES CONSULTÉES

- Allen, R. B. (2005). Using Information Visualization to Support Access to Archival Records. *Journal of Archival Organization*, 3(1), 37-49. doi.org/10.1300/J201v03n01_04
- Bahde, A. (2017). Conceptual Data Visualization in Archival Finding Aids: Preliminary User Responses. *portal: Libraries and the Academy*, 17(3), 485-506. doi.org/10.1353/pla.2017.0031
- Berry, D. M. (dir.). (2012). *Understanding Digital Humanities*. Basingstoke, Royaume-Uni: Palgrave Macmillan.
- Burdick, A., Drucker, J., Lunenfeld, P., Presner, T. et Schnapp, J. (2012). *Digital Humanities*. Cambridge, MA: MIT Press.
- Burnard, L. (2012). Du *literary and linguistic computing* aux *digital humanities*: retour sur 40 ans de relations entre sciences humaines et informatique. Dans P. Mounier (dir.), *Read/Write book 2* (p. 45-58). Marseille, France: OpenEdition Press.
- Busa, R. (1980). The Annals of Humanities Computing: The Index Thomisticus. *Computers and the Humanities*, 14(2), 83-90. Repéré à www.alice.id.tue.nl/references/busa-1980.pdf
- Bush, V. (1945). As We May Think. *The Atlantic Monthly*, 176(1), 101-108. Repéré à www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/
- Crane, G. R. et Wulfman, C. E. (2003, mai). *Towards a Cultural Heritage Digital Library*. Communication présentée à la 3^e conférence commune ACM/IEEE sur les bibliothèques numériques, Houston, Texas. doi.org/10.1109/JCDL.2003.1204845
- Dalbello, M. (2011). A Genealogy of Digital Humanities. *Journal of Documentation*, 67(3), 480-506. doi.org/10.1108/00220411111124550
- Gibbs, F. (2011). Critical Discourse in Digital Humanities. *Journal of Digital Humanities*, 1(1). Repéré à journalofdigitalhumanities.org/1-1/critical-discourse-in-digital-humanities-by-fred-gibbs/
- Granjon, F. (2016). Présentation du dossier: les sciences humaines et sociales au prisme du *digital turn*. *Variations*, 19(2016). doi.org/10.4000/variations.726
- Hockey, S. (2004). The History of Humanities Computing. Dans S. Schreibman, R. Siemens et J. Unsworth (dir.), *A Companion to Digital Humanities* (1^{re} éd., p. 13-19). Oxford, Royaume-Uni: Wiley-Blackwell. doi.org/10.1002/9780470999875.ch1

- Klein, J. T. (2015). *Interdisciplining Digital Humanities: Boundary Work in an Emerging Field*. Ann Arbor, MI: University of Michigan Press.
- Latham, S. (2004). "New age" scholarship: The Work of Criticism in the Age of Digital Reproduction. *New Literary History*, 35(3), 411-426. doi.org/10.1353/nlh.2004.0043
- Lyman, P. (1984). Reading, Writing and Word Processing: Towards a Phenomenology of the Computer Age. *Qualitative Sociology*, 7(1-2), 75-89. Repéré à link.springer.com/article/10.1007/BF00987108
- Marchionini, G. (2000). Evaluating Digital Libraries: A Longitudinal and Multifaceted View. *Library Trends*, 49(2), 304-333. Repéré à core.ac.uk/download/pdf/4817672.pdf
- McLuhan, M. (1968). *Pour comprendre les médias: les prolongements techniques de l'homme* (traduit par J. Paré). Montréal, Québec: Éditions Hurtubise HMH.
- Nelson, T. (1965, août). *A File Structure for the Complex, the Changing and the Intermediate*. Communication présentée à la 20^e conférence de l'Association for Computing Machinery (ACM), Cleveland, Ohio. Repéré à archive.org/details/nelson-file-structure
- Poole, A. H. (2017). The Conceptual Ecology of Digital Humanities. *Journal of Documentation*, 73(1), 91-122. Repéré à www.researchgate.net/publication/311921946_The_Conceptual_Ecology_of_Digital_Humanities
- Poole, A. H. et Garwood, D. A. (2018). "Natural allies": Librarians, Archivists, and Big Data in International Digital Humanities Project Work. *Journal of Documentation*, 74(4), 804-826. doi.org/10.1108/JD-10-2017-0137
- Reed, A. (2014). Managing an Established Digital Humanities Project: Principles and Practices from the Twentieth Year of the William Blake Archive. *Digital Humanities Quarterly*, 8(1). Repéré à www.digitalhumanities.org/dhq/vol/8/1/000174/000174.html
- Robinson, L., Priego, E. et Bawden, D. (2015, mai). *Library and Information Science and Digital Humanities: Two Disciplines, Joint Future?* Communication présentée au 14^e symposium international sur les sciences de l'information, Zadar, Croatie. Repéré à openaccess.city.ac.uk/id/eprint/11889
- Ruhleder, K. (1995). Reconstructing Artifacts, Reconstructing Work: From Textual Edition to On-Line Databank. *Science, Technology, & Human Values* 20(1), 39-64. Repéré à jstor.org/stable/689880?seq=1#page_scan_tab_contents
- Sinatra, M. E. et Vitali-Rosati, M. (2014). Histoire des humanités numériques. Dans M. E. Sinatra et M. Vitali-Rosati (dir.), *Les pratiques de l'édition numérique* (p. 49-62). Montréal, Québec: Presses de l'Université de Montréal.
- Thomas III, W. G. (2016). The Promise of Digital Humanities and the Contested Nature of Digital Scholarship. Dans S. Schreibman, R. Siemens et J. Unsworth (dir.), *A New Companion to Digital Humanities* (2^e éd., p. 524-537). Oxford, Royaume-Uni: Wiley-Blackwell.
- Whitelaw, M. (2009). Visualising Archival Collections: The Visible Archive Project. *Archives and Manuscripts*, 37(2), 22-40. Repéré à mtchl.net/assets/VisualisingArchivalCollections.pdf