



HAL
open science

En finir avec les Humanités Numériques : Penser les SHS avec les algorithmes

Christian Henriot, Cécile Armand

► To cite this version:

Christian Henriot, Cécile Armand. En finir avec les Humanités Numériques : Penser les SHS avec les algorithmes. Beyond Digital Humanities: How computational methods are reshaping scholarly research, Cécile Armand, Sep 2023, Aix en Provence, France. halshs-04254636

HAL Id: halshs-04254636

<https://shs.hal.science/halshs-04254636>

Submitted on 23 Oct 2023

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

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



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

En finir avec les Humanités Numériques : Penser les SHS avec les algorithmes

Cécile Armand (Aix-Marseille University)
Christian Henriot (Aix-Marseille University)

« L'historien de demain sera programmeur ou il ne sera pas »
Emmanuel Leroy-Ladurie, *Le Nouvel Observateur*, 8 mai 1968

Introduction

Les sciences humaines et sociales (SHS) forment un ensemble de disciplines à géométrie variable, incluant généralement la littérature, les langues, l'histoire, la géographie, la sociologie et l'anthropologie. Historiquement, ces disciplines sont ancrées dans l'analyse textuelle, les méthodologies interprétatives et un engagement profond dans la compréhension de l'expérience humaine. Au cours des dernières décennies, ce paysage a été bouleversé par l'avènement des Humanités Numériques (HN)—un domaine qui promettait de marier les traditions de la recherche humaniste avec le dynamisme des technologies numériques. À première vue, les HN ont émergé comme un phare de l'innovation, offrant des outils et des techniques qui assuraient de révolutionner la recherche. D'un autre côté, les esprits sceptiques ont pointé le caractère oxymorique de l'expression, soulignant l'antagonisme irréductible de la tradition humaniste et des sciences informatiques.

Comme souvent avec les nouveaux paradigmes, l'euphorie initiale apporte avec elle un ensemble de défis. L'un des défis principaux est d'ordre définitionnel. Que signifie précisément "Humanités Numériques" ? Le sens du terme s'est élargi au point de perdre en substance et en spécificité. La numérisation d'un manuscrit médiéval relève-t-elle des HN de la même manière qu'un algorithme analysant des motifs à travers des milliers de textes littéraires ? Il existe une littérature abondante s'attendant à définir les Humanités Numériques (HN) (cf. références). Cet article ne prétend pas fournir une nouvelle définition. Au contraire, il propose d'abandonner les perspectives définitionnelles au profit d'un cadrage plus heuristique et plus constructif du débat. La très grande polyvalence des HN, qui fait leur force, est également à l'origine de leur ambiguïté. L'un des principaux dangers des HN est d'avoir involontairement enfermé la recherche en sciences humaines dans un ensemble de dichotomies de type ancien/nouveau, traditionnel/numérique, quantitatif/qualitatif ou lecture rapprochée/lecture distante. Loin de transcender ces clivages, l'affirmation des humanités numériques n'a fait que les renforcer.

Cet essai aborde les humanités numériques dans une perspective critique, celle de deux historiens ancrés dans la recherche empirique. Il vise à interroger le concept général et parfois nébuleux de « numérique » dans les sciences humaines. En particulier, il invite à se détacher des étiquettes englobantes pour se concentrer sur le potentiel transformateur des **méthodes computationnelles à l'intérieur même des différentes disciplines**. L'objectif n'est pas de rejeter les Humanités Numériques (HN) ou notions associées, mais de repartir de notre pratique disciplinaire pour faire en sorte que le tournant numérique soit autre chose qu'un *buzzword*,

une promesse ou une phobie. L'enjeu ultime est d'enraciner **lesdites humanités numériques dans la rigueur méthodologique, les sources, les questionnements et les besoins propres à chaque discipline** et dans les valeurs fondamentales des sciences humaines et sociales.

Après avoir souligné les limites de l'étiquette HN, nous donnerons un aperçu des **méthodes computationnelles** existantes en essayant de rendre compte de leurs nuances. Les deux dernières parties plaident pour que ces **méthodes ne soient pas conçues comme de simples appendices, mais soient pleinement intégrées aux pratiques des chercheurs en sciences humaines et sociales.**

Les Limites de l'étiquette « Humanités Numériques »

Les **Humanités Numériques** (HN) ont souvent été présentées comme une nouvelle frontière des sciences humaines, promettant de redéfinir en profondeur les contours de la recherche académique. Cependant, à mesure qu'il s'est popularisé, le terme est devenu de plus en plus insaisissable. La difficulté est accentuée en France par le fait que l'expression « humanités numériques » a été adoptée sans distance critique par les chercheurs et surtout par les décideurs (ministères, agences de financement, etc.), à la différence des milieux académiques anglo-saxons, où le champ s'est développé de manière progressive suivant une logique *bottom up*.

L'ambiguïté du terme, tout en reflétant son étendue et la variété de ses applications, pose un certain nombre de problèmes qui méritent un examen plus approfondi. Tout d'abord, le terme « numérique » est **large et comporte de multiples facettes**. Alors que le numérique imprègne aujourd'hui presque tous les aspects de nos vies, qu'est-ce que cela signifie pour les sciences humaines d'être « numériques » ? S'agit-il de numériser des archives ou des documents imprimés, de traiter des données par des algorithmes, ou encore de diffuser en ligne les résultats de la recherche ? Le terme HN, dans son effort pour englober toutes ces dimensions, est devenu une expression **fourre-tout** qui aplatit les nuances et dilue la spécificité des pratiques qu'il englobe. Victime d'un effet de mode, le terme s'est imposé comme une sorte de passage obligé que l'on distille dans les plaquettes d'enseignement universitaires, les demandes de subventions ou de postes et dans les *curriculum vitae* des candidats. Avec la banalisation des humanités numériques, il est devenu difficile de différencier les véritables praticiens de ceux qui emploient simplement le terme pour se conformer aux attentes du moment.

Ceci débouche sur un deuxième problème : le **risque de simplification** excessive. En regroupant sous une même bannière un éventail de pratiques extrêmement diversifiées, l'expression « humanités numériques » dissout les spécificités de chacune. Par exemple, la numérisation de textes anciens, qui a sa propre valeur intrinsèque, est une entreprise très différente de l'utilisation du *machine learning* pour modéliser des phénomènes linguistiques ou des genres littéraires dans de vastes corpus à travers plusieurs siècles. Placer ces tâches très différentes sous la même bannière HN revient à faire l'impasse sur les enjeux méthodologiques, épistémologiques et pratiques spécifiques qui les animent.

Plus largement, l'expression HN peut alimenter une forme de **techno-utopie universaliste qui consiste à adopter en bloc l'innovation technologique sans** se poser la question de sa pertinence pour une recherche donnée, de ses limites et ses biais éventuels. Tous les outils numériques ne conviennent pas nécessairement à toutes les questions. Les appliquer sans discernement peut conduire à des analyses superficielles voire erronées. Inversement, le techno-utopisme éveille souvent le **scepticisme, la méfiance** voire la résistance de nombreux

membres de la communauté académique. Ceux qui osent s'aventurer dans le territoire numérique peuvent devenir aux yeux de leurs pairs des hérétiques ayant renié leur discipline et succombé au démon de la culture *geek*. Ces perceptions persisteront tant que l'on continuera à considérer le tournant numérique comme une mutation purement technique. Or, l'obstacle principal pour les chercheurs en sciences humaines n'est pas technologique, mais bien plus profondément herméneutique. La question fondamentale n'est pas de déterminer ce que sont les humanités numériques ou si le numérique est compatible avec les humanités, mais plutôt de se demander comment nous pouvons nous approprier des outils qui n'ont pas été conçus pour, et encore moins par, des chercheurs en sciences humaines. Comment pouvons-nous interpréter les résultats produits par les algorithmes pour construire un récit cohérent et nuancé redonnant toute sa place à l'expérience humaine, qui reste notre objet d'étude inconditionnel ?

Enfin, l'étiquette englobante des HN peut avoir de sérieuses **implications institutionnelles**. Si les institutions européennes ont été plus lentes à réagir par rapport aux institutions anglo-saxonnes et maintenant chinoises, le mouvement est désormais indéniable dans les économies dites « développées ». Depuis que les universités, les ministères et les agences de financement se sont lancées à corps perdus dans la course au numérique, il existe cependant un risque que les ressources soient canalisées vers des projets « numériques » en apparence, au détriment de propositions qui offrent une véritable innovation méthodologique sans nécessairement agiter la bannière numérique.

L'essor des humanités numériques s'est accompagné de toutes sortes de préjugés qui trahissent la relation ambivalente que les chercheurs en sciences humaines entretiennent avec les ordinateurs depuis l'expérience traumatisante de la Cliométrie dans les années 1960. Ces perceptions ont largement contribué à marginaliser les humanités numériques dans la recherche académique. Voici quelques-uns des clichés les plus répandus :

1. **La querelle des Anciens et des Modernes** repose sur l'idée qu'il existerait deux espèces bien distinctes de chercheurs : ceux qui auraient une vocation pour les outils numériques, d'un côté, et ceux qui seraient irrémédiablement voués aux méthodes « traditionnelles », de l'autre.
2. **Le biais générationnel** est une déclinaison du précédent. Du fait de leur jeunesse, les jeunes chercheurs « biberonnés » à la culture numérique auraient des dispositions particulières pour les outils numériques, tandis que les plus anciens nés à l'ère du papier ne pourraient jamais combler cet écart générationnel.
3. **La magie du numérique et le paradigme révolutionnaire** : Les Humanités numériques nourrissent bon nombre de fantasmes auprès des non-praticiens, qui attribuent parfois une sorte de pouvoir magique aux outils numériques, imaginant que ces méthodes vont instantanément révolutionner l'état de l'art.
4. « **Rien de nouveau sous le soleil** » : Ces grandes espérances laissent bientôt place à une déception proportionnelle : « Mais au fond, qu'avez-vous démontré que l'on ne savait déjà ? Pourquoi tant d'efforts pour un si maigre résultat ? » Telles sont les objections auxquelles les praticiens qui tentent de démontrer la validité de telles méthodes sont bien souvent confrontés.
5. « **Ne jetons pas le bébé avec l'eau du bain !** » : D'autres chercheurs, sans être foncièrement hostiles au numérique, consacrent beaucoup d'énergie à en disséquer les biais et les difficultés. Certains mettent en lumière des préoccupations éthiques, comme l'accès inégal aux plateformes numériques à l'échelle mondiale. D'autres pointent les lacunes des algorithmes, comme les imperfections de la reconnaissance optique de caractères sur des textes originaux de piètre qualité.

Si une approche lucide du domaine numérique est absolument essentielle, ces critiques exagérées, infondées ou mal informées nuisent à une compréhension raisonnée de ce que les algorithmes peuvent (et ne peuvent pas) apporter aux sciences humaines.

L'expression « Humanités Numériques » possède néanmoins certaines vertus : elle traduit le dynamisme de la recherche en sciences humaines et indique une progression vers l'interdisciplinarité. Il est cependant crucial d'aborder le domaine numérique avec discernement. Face à l'arrivée des nouveaux modèles génératifs de type ChatGPT, les chercheurs doivent plus que jamais privilégier la précision et la clarté. Depuis l'apparition des *humanities computing* qui les ont précédées, les humanités numériques n'ont cessé d'évoluer et ont donné naissance à divers sous-domaines. On peut citer par exemple l'histoire numérique, apparue en 1993, qui constitue l'une des premières formulations associant le qualificatif numérique à une discipline spécifique. Dans les années 2010, la sociologie numérique et l'anthropologie numérique ont à leur tour émergé comme des champs autonomes. Dans d'autres cas, ce sont des méthodologies spécifiques, telles que les Systèmes d'Information Géographique (SIG), qui ont catalysé la formation de nouveaux domaines, comme les humanités spatiales (*spatial humanities*).

Plus récemment, l'usage du terme « computationnel » a commencé à s'imposer dans certaines disciplines. S'il n'y a pas eu, à notre connaissance, de « science politique numérique » ou d'« économie numérique » (au sens de sciences économiques), depuis les années 2000, les **sciences sociales computationnelles** (*computational social sciences*) ont émergé comme le pendant des HN dans ces disciplines, donnant naissance à des communautés, des institutions, des conférences et des revues distinctes. Par la suite, le terme computationnel s'est également étendu aux sciences humaines, aboutissant à la formation d'un nouveau champ, celui des humanités computationnelles (*computational humanities*). Ce glissement exprime un engagement plus poussé avec des méthodes computationnelles avancées, telles que l'apprentissage automatique ou les simulations, et une immersion plus profonde dans le domaine des algorithmes (Tableau 1). Ce nouveau concept englobant présente cependant les mêmes risques que les précédents.

Tableau 1. Mise en perspective des concepts liés au numérique

| Caractéristiques | Digital Humanities (DH) | Computational Social Science (CSS) | Computational Humanities (CH) |
|------------------------|---|---|---|
| Disciplines concernées | Toutes les disciplines des sciences humaines et sociales. | Disciplines utilisant prioritairement des données non textuelles (sociologie, économie) | Disciplines centrées sur les textes (histoire, langues, littérature...) |

| | | |
|--------------------------|--|---|
| Objectifs | Numériser, analyser et visualiser des données ; enrichir la recherche « traditionnelle » avec des outils numériques. | Comprendre, modéliser ou prédire des phénomènes sociaux, culturels ou linguistiques à l'aide de méthodes computationnelles, en s'appuyant sur des données empiriques. |
| Orientations / pratiques | Large spectre de pratiques allant de la numérisation de documents à la construction de bases de données, en passant par l'analyse et la visualisation de données au moyen de logiciels prêt-à-l'emploi. Différents degrés d'implication et de compétence de la part des chercheurs. | Engagement profond avec les algorithmes nécessitant la maîtrise d'un langage de programmation, qui permet au chercheur de travailler dans un environnement intégré et d'acquérir une large autonomie. |
| Interdisciplinarité | Dialogue minimaliste avec des ingénieurs et spécialistes des données. | Dialogue soutenu et régulier entre chercheurs en informatique et SHS, nécessitant une immersion profonde dans l'autre discipline pour en comprendre les concepts. |

Les méthodes computationnelles : une frontière encore peu explorée dans les sciences humaines

Parce qu'il évoque un potentiel d'innovation, le terme numérique captive l'imagination. Cependant, si l'on creuse un peu sous la surface, il apparaît très vite que le véritable pouvoir de transformation ne réside pas dans les objets ou les technologies numériques, mais dans le monde invisible des algorithmes qui les sous-tendent et dans les méthodes computationnelles qui les prolongent. La notion de « computationnel » appliquée aux humanités renvoie à une utilisation poussée des algorithmes pour analyser les données des sciences humaines, qui transcende la simple numérisation des documents ou l'utilisation superficielle d'un logiciel

prêt-à-l'emploi. Alors que le numérique désigne tantôt un outil, un support ou un format (le résultat d'une numérisation), le computationnel signale sans ambiguïté une heuristique : une démarche rigoureuse, systématique et reproductible qui emprunte avec discernement ses concepts et ses outils aux sciences informatiques.

Considérons par exemple le domaine littéraire. Alors que les bases de données et les interfaces numériques permettent d'accéder à de vastes bibliothèques en ligne et d'en consulter les contenus, les méthodes computationnelles permettent d'identifier des thèmes latents, des motifs, des régularités ou des anomalies impossibles à discerner au moyen d'une lecture linéaire traditionnelle. Ces méthodes de *distant reading* sont aussi employées pour analyser les genres, les styles et les phénomènes d'intertextualité à travers de vastes corpus. De même, en histoire, alors que la numérisation des archives améliore leur accessibilité, ce sont les méthodologies computationnelles comme le traitement automatique du langage (TAL) qui permettent d'en extraire l'information historique pertinente, tandis que les SIG permettent de cartographier les lieux historiques, de reconstituer la géographie des événements et d'analyser les processus spatiaux qui s'y déroulent (migrations, expansion urbaine). Les SIG peuvent également être appliqués aux textes littéraires, par exemple pour cartographier les déplacements des personnages ou les émotions qu'ils attachent aux lieux. Autrement dit, les méthodes computationnelles ne viennent pas simplement s'ajouter à la boîte à outils existante ; elles redéfinissent en profondeur les questions que les chercheurs peuvent poser et les connaissances qu'ils peuvent en tirer.

Même si les sources visuelles, sonores ou matérielles constituent des matériaux précieux et parfois irremplaçables, **le texte reste central dans la recherche en sciences humaines**. Tant que les textes numérisés n'étaient disponibles que sous la forme d'images scannées à partir des documents originaux, les chercheurs en restaient à un usage purement consultatif. Ils continuaient à lire les textes numérisés comme ils lisaient leurs originaux imprimés, avec pour seule différence que ces textes se trouvaient désormais sous une forme dématérialisée. Le vrai changement est intervenu à partir du moment où la reconnaissance optique de caractère (OCR) a permis de transformer les images en données textuelles (*full text*). À partir de là, les chercheurs se sont rapidement retrouvés submergés par un océan de textes en continuelle expansion, dépassant de loin les capacités de l'esprit humain. D'un autre côté, le texte ainsi transformé en données devint traitable par les algorithmes, offrant alors des possibilités d'exploration et d'analyse jusqu'ici inenvisageables. Les chercheurs en littérature ont été les premiers à faire face au déluge des *big data*, notamment parce que les œuvres littéraires publiées avant le 20^e siècle relevaient du domaine public.

Les méthodes computationnelles sont le produit de développements qui se déroulent sur plusieurs décennies. Jusqu'aux années 2000, elles n'ont eu qu'un impact limité sur les sciences humaines et sociales, restant cantonnées à des niches comme celles du *humanities computing*. Cependant, l'essor d'Internet, l'augmentation exponentielle de la puissance de calcul, le développement fulgurant des méthodes d'apprentissage automatique (du *machine* au *deep learning*) et la généralisation de l'intelligence artificielle ont en quelques années transformé en profondeur le paysage, si bien que les chercheurs en sciences humaines se sont retrouvés face à un outillage radicalement nouveau auxquels ils n'étaient pas préparés.

L'adoption des méthodes computationnelles ne va pas de soi. Ces méthodes **requièrent une littératie particulière, à laquelle les humanistes de formation sont étrangers** : une certaine capacité à penser de manière algorithmique, une bonne compréhension des structures de données et une solide connaissance des notions statistiques. Formés à la lecture rapprochée des

textes, privilégiant les analyses qualitatives et généralement dépourvus de culture mathématique, les chercheurs en sciences humaines sont mal armés. Beaucoup devront d'abord surmonter leur réticence à sortir de leur zone de confort, accepter de prendre le temps de lire et se former à des concepts éloignés de leur domaines d'expertise bien circonscrits. Force est d'admettre qu'il y a une courbe d'apprentissage, qui peut sembler raide au début, avant d'acquérir les compétences nécessaires au bon usage des méthodes computationnelles. Cette littératie est essentielle pour pouvoir identifier les méthodes qui correspondent le mieux aux objectifs de recherche que l'on s'est fixés et s'assurer que leur application reste conforme à l'éthique humaniste.

L'adoption des méthodes computationnelles ne doit pas se faire de manière uniforme. Elle doit être adaptée aux besoins de chaque discipline et même de chaque recherche individuelle. Si le tournant numérique a représenté une révolution pour les humanistes, l'avenir de la recherche en science humaines repose sur une appropriation nuancée et sélective des méthodes computationnelles.

L'introduction des méthodologies computationnelles dans les sciences humaines nécessite **d'élargir la formation des humanistes et de renforcer les pratiques d'interdisciplinarité**. La complexité des algorithmes, du *machine learning* et de l'intelligence artificielle échappent et continueront d'échapper à la plupart des chercheurs en sciences humaines. À l'inverse, si les chercheurs en informatique sont les mieux placés pour créer des algorithmes sophistiqués, il leur manque les compétences herméneutiques nécessaires pour analyser, contextualiser et interpréter la portée sociale, culturelle ou historique des données que ces algorithmes traitent.

Le tournant computationnel exige de la part des chercheurs un véritable effort pour s'engager en direction de l'autre discipline. Ces nouvelles synergies entre l'informatique et les sciences humaines peuvent se traduire par la co-crédation d'outils et de méthodes à la fois technologiquement très avancés mais toujours ancrés dans les fondements de la pensée humaniste. Un exemple de collaboration en ce sens est l'application [HistText](#) — une librairie R permettant d'exploiter de vastes corpus de textes historiques et multilingues — développée par une équipe d'historiens, d'informaticiens et de *data scientists* qui ont travaillé ensemble de manière rapprochée pendant plus de cinq ans (Blouin et al., 2023). Une telle collaboration interdisciplinaire garantit que les algorithmes sont adaptés aux questions spécifiques des chercheurs, tandis que les données produites sont interprétées avec la profondeur et la nuance que seule une formation humaniste peut apporter.

L'avenir des méthodes computationnelles réside dans cet esprit de collaboration. À mesure que les données deviennent une réalité inéluctable de la recherche comme de la vie sociale, **les frontières entre les disciplines devront s'assouplir**. L'une des clés pour réussir cette transition est de créer des espaces, à l'image du [Center for Interdisciplinary Digital Research \(CIDR\)](#) à Stanford, où les chercheurs en informatique, en sciences des données et en sciences humaines et sociales, peuvent se retrouver à tout moment pour échanger, apprendre les uns des autres, se poser des défis mutuels et progresser ensemble vers une compréhension à la fois plus fine et plus holistique du monde.

Le tournant critique : de l'utilisation à la compréhension des outils

L'attrait de la technologie réside généralement dans sa promesse d'efficacité, sa capacité à automatiser des tâches et à révéler des motifs (patterns) invisibles à l'œil nu. Dans le paysage des Humanités Numériques (HN) ou Computationnelles (HC), cet attrait est palpable. Les outils numériques offrent des fonctionnalités alléchantes pour analyser, visualiser et diffuser des données, tout est restant relativement accessibles aux chercheurs qui n'ont aucune compétence en programmation. Les fournisseurs commerciaux multiplient les promesses révolutionnaires et les visualisations d'un esthétisme éblouissant, mais pas toujours efficace pour la recherche en sciences humaines. L'offre est devenue vertigineuse et il est très difficile de se repérer. Les chercheurs doivent cependant renoncer à l'illusion d'un outil à tout faire et se méfier des logiciels prêts à l'emploi qui n'ont souvent pas été conçus pour la recherche.

Les humanités numériques sont arrivées à un tournant critique. Le temps de la fascination est derrière nous. Le moment est venu de dépasser l'utilisation passive des outils pour parvenir à une compréhension plus profonde des algorithmes qui les sous-tendent afin d'en démasquer les biais potentiels.

Les sciences humaines ont une longue tradition d'enquête critique — une capacité à interroger, interpréter et déconstruire les textes et les données que les chercheurs ont entre les mains. Cette capacité critique fait leur force. Ils doivent apprendre à l'appliquer au domaine numérique comme ils ont appris à le faire sur les textes et autres sources classiques. À mesure qu'elles deviennent partie intégrante de notre environnement, les techniques numériques et computationnelles doivent impérativement faire l'objet d'une appropriation critique. Cela implique de se demander non seulement ce que les outils peuvent faire, mais plus fondamentalement comment ils fonctionnent, quels postulats les sous-tendent, quels biais ils peuvent introduire, et quelles sont les implications épistémologiques de leur utilisation. Cela implique aussi d'expérimenter, de tester différentes hypothèses et comparer différents algorithmes avant de choisir celui qui convient le mieux aux questions et aux données. En somme, le tournant computationnel pousse les sciences humaines à se rapprocher davantage des sciences expérimentales. Les historiens et les littéraires, notamment, doivent aussi se demander ce qu'est une « source » ou un « corpus » numérique et qu'est-ce que cela implique de travailler à partir de l'avatar numérique d'une source ou d'un corpus. Répondre correctement à ces questions nécessite à la fois une compréhension profonde du raisonnement computationnel ainsi que des compétences herméneutiques solidement ancrées dans la tradition humaniste.

Considérons, par exemple, un travail d'analyse de texte qui utiliserait un algorithme de classification pour catégoriser les genres littéraires. L'algorithme parviendra toujours à réaliser la tâche demandée, mais le résultat ne correspondra pas forcément aux exigences du chercheur. Face aux classifications produites de manière automatique, un humaniste se doit d'interroger les critères de catégorisation, les biais potentiels induits par les données d'apprentissage et les implications plus larges de la détermination algorithmique des genres. Cette démarche critique garantit que les méthodes computationnelles ne diluent pas mais renforcent au contraire la profondeur analytique de la recherche en sciences humaines. En somme, il ne s'agit jamais d'utiliser passivement des outils ; il s'agit de les comprendre et d'en corriger les biais pour s'assurer qu'ils sont en adéquation avec les besoins épistémologiques des sciences humaines.

Ce tournant critique est essentiel pour des raisons non seulement épistémologiques, mais aussi **éthiques**. C'est aujourd'hui un fait reconnu que les algorithmes perpétuent des biais fondés sur le genre, la race ou la classe sociale. Étant donné le rôle croissant que jouent les algorithmes dans nos sociétés, les chercheurs ont la responsabilité de s'assurer que leurs méthodes respectent les valeurs fondamentales d'équité, d'inclusion et de rigueur académique. Cela implique non seulement de dépasser une utilisation superficielle des outils, mais aussi de s'immerger profondément dans leur mécanique, leurs données d'apprentissage et leurs écueils potentiels. Parvenir à cette appropriation critique exige un véritable **dialogue interdisciplinaire**, qui nécessite des deux côtés une formation à l'autre discipline, ou du moins une compréhension minimale de ses concepts fondamentaux. Ce n'est qu'en comprenant le langage computationnel que les chercheurs en sciences humaines peuvent s'engager efficacement dans l'analyse de leurs données et construire avec les sciences informatiques une interdisciplinarité forte fondée sur la compréhension mutuelle.

En d'autres termes, l'herméneutique doit être la boussole guidant notre voyage numérique. Il ne suffit pas de manier des outils numériques ; nous devons les disséquer pour mieux les comprendre et les situer de manière critique dans le spectre plus large de nos méthodes. C'est la seule façon de garantir que le tournant numérique dans les sciences humaines ne sera pas seulement un tournant technologique, mais aussi un tournant épistémologique ancré dans la rigueur et les valeurs éthiques qui animent depuis toujours les disciplines humanistes.

Transformations spécifiques aux disciplines

Pour mieux apprécier le potentiel des méthodes computationnelles en sciences humaines, il est utile d'examiner les différentes disciplines au cas par cas pour comprendre comment chacune est transformée de manière spécifique. Trois cas retiendront ici notre attention : la littérature, l'histoire et la sociologie.

- La littérature : dépasser le clivage entre *close* et *distant reading*

Traditionnellement, les études littéraires ont été ancrées dans le *close reading* — une lecture linéaire et méticuleuse des textes, attentive à leurs nuances et leurs particularités stylistiques. Cependant, avec la numérisation massive des textes littéraires et la constitution de vastes corpus accessibles en ligne, le *close reading* se heurte inévitablement à un problème d'échelle. Pour contourner l'obstacle, des chercheurs comme Franco Moretti, Matthew Jockers ou plus récemment Ted Underwood ont introduit les concepts de *distant reading* et de macro-analyse. Cette approche s'appuie sur des algorithmes pour analyser de vastes corpus de textes et identifier des motifs et des tendances sur plusieurs siècles. Elle est utilisée par exemple pour discerner les évolutions stylistiques, l'émergence des genres littéraires, les changements thématiques et les influences littéraires à une échelle auparavant inimaginable. Cette approche a fait l'objet de nombreuses critiques, notamment celle de réduire l'analyse littéraire à un comptage statistique des mots, de diluer la singularité des auteurs et des styles, ou encore d'ignorer le contexte de production des oeuvres. Cependant, le développement des techniques d'apprentissage profond (*deep learning* et *transformers*) et les avancées récentes dans le domaine du traitement automatique du langage neutralisent la plupart de ces critiques. Des méthodes d'analyse comme le *topic modeling* (la détection de thème latent dans un corpus à partir des co-occurrences les plus fréquentes) ou les *word embeddings* (vecteurs de mots projetés dans un espace multidimensionnel) permettent désormais de mieux restituer le

contexte sémantique et de naviguer sur un continuum entre différentes échelles de lecture. Ce faisant, l'opposition binaire entre *distant* et *close reading* apparaît de moins en moins pertinente.

- **L'histoire : des archives numériques aux analyses computationnelles**

Longtemps associée à la poussière des archives et des manuscrits, la recherche historique n'a pas échappé à la révolution numérique. Si la numérisation des archives a contribué à « démocratiser » l'accès aux documents historiques, ce sont les méthodes computationnelles qui ont véritablement transformé les pratiques historiographiques. Par exemple, la reconnaissance d'entités nommées, une technique de traitement automatique du langage qui consiste à détecter les noms de personnes, d'organisations, de lieux, et d'événements, permet d'extraire avec précision de l'information historique à travers de vastes corpus de textes. Les méthodes d'analyse textuelle peuvent être utilisées pour analyser les pratiques discursives dans une perspective comparative et de longue durée, une démarche qui répond à l'exigence fondamentale pour l'historien de croiser et contextualiser ses sources. Des techniques comme les Systèmes d'Information Géographique (SIG) permettent de cartographier la géographie des événements historiques, d'analyser les phénomènes spatiaux comme les migrations à différentes échelles et de les interpréter en leur superposant différentes couches de données géo-historiques. Les outils d'analyse de réseau aident à mieux saisir la complexité des relations sociales, des liens diplomatiques ou des échanges commerciaux. Les méthodes computationnelles offrent ainsi de multiples voies d'accès vers le passé, que l'historien peut combiner pour construire un récit plus holistique.

- **La sociologie et l'anthropologie : des enquêtes de terrain à l'ethnographie numérique**

Les sociologues et les anthropologues s'intéressent aux structures et aux dynamiques sociales. Les enquêtes de terrain et les entretiens individuels ont longtemps constitué une méthode privilégiée. A l'heure où des millions d'individus à travers le monde laissent à chaque instant un nombre infini de traces numériques sur de multiples plateformes, les méthodes traditionnelles d'enquête se heurtent à un problème d'échelle. Les outils numériques offrent différentes solutions pour appréhender cette mutation. Outre les outils de fouille de données et de *web scraping* qui permettent de collecter les données, l'analyse de réseau, peut là encore fournir un éclairage sur la structure et la dynamique des liens sociaux, les rapports de pouvoir et de hiérarchie, ou la formation de communautés. L'ethnographie numérique peut tirer profit des plateformes multimédias, des réseaux sociaux et des forums en ligne, pour étudier les phénomènes d'opinion ou de rumeurs. Grouillant d'interactions, de débats et d'échanges, mais aussi de biais et d'asymétries, ces espaces numériques s'affirment comme les nouveaux « terrains » de la sociologie. Plus complexes à appréhender que les enquêtes dont les chercheurs construisent le cadre et maîtrisent le déroulé, ces sites offrent cependant des perspectives en temps réel sur les changements sociétaux, les tendances culturelles et les phénomènes sociaux émergents, que les sociologues peuvent saisir au moyen de méthodes computationnelles appropriées, sans pour autant négliger le terrain plus familier de la vie sociale « hors ligne ».

Ces exemples mettent en lumière un fait crucial : la transformation impulsée par les méthodes computationnelles ne se traduit pas simplement par l'ajout de nouveaux outils, mais par une **refonte fondamentale des pratiques disciplinaires**. Dans chaque discipline, le tournant computationnel aboutit à neutraliser les dichotomies de type quantitatif/qualitatif ou *close /*

distant reading au profit d'une approche multi-scalaire qui rend possible l'exploration de vastes ensembles de données sans perdre en précision ni en complexité. Par exemple, le *topic modeling* permet de naviguer du mot au corpus dans son ensemble, en passant par la phrase, le document ou le *topic* commun à plusieurs documents. Dans le même esprit, l'analyse de réseau permet de naviguer des nœuds individuels à la structure globale du réseau, en passant par les dyades ou les communautés qui le constituent. Les chercheurs qui s'appuient sur ces méthodes ont intérêt à compléter les visualisations qu'elles fournissent par un examen minutieux des mesures statistiques qui les sous-tendent. Une solide connaissance du contexte reste indispensable pour interpréter correctement la situation particulière de chaque individu (ou document) au sein de la structure (ou corpus) et donner sens aux communautés (ou *topics*) qui s'y forment. C'est à partir de ces transformations spécifiques, ancrées dans les concepts et les pratiques propres à chaque discipline, que se dessine l'avenir computationnel de la recherche en sciences humaines.

L'avenir : intégration plutôt que substitution

La généralisation des algorithmes et de l'intelligence artificielle ouvre un débat crucial concernant leur impact sur la recherche et sur l'avenir des sciences humaines. Les méthodes computationnelles vont-elles remplacer les méthodes « traditionnelles » ? Ou bien viendront-elles simplement suppléer l'outillage existant ? Comme souvent, la réalité se situe quelque part entre les deux, dans le continuum reliant les méthodes bien établies à l'adoption progressive de technologies encore expérimentales. Les méthodes computationnelles ne vont pas faire disparaître les méthodes établies. Elles doivent fusionner avec elles de manière à renforcer la boîte à outils des humanistes.

Le souci de comprendre l'expérience humaine dans toute sa complexité et de lui donner sens à travers l'histoire et les sociétés est au cœur des sciences humaines. Quel que soit l'outillage employé, ce noyau fondamental reste inchangé. Les méthodes que les sciences humaines et sociales ont élaborées sont façonnées par la rigueur, la nuance et la souplesse du cadre interprétatif. C'est sur ce socle fondateur que les nouvelles méthodes doivent être implantées.

Malgré leur potentiel transformateur, les méthodes computationnelles ne sont pas des panacées. Elles offrent des fonctionnalités spécifiques — les variations d'échelle, la détection de motifs macroscopiques, ou encore la possibilité de fonder les interprétations sur des données systématiques. Cependant, elles ont aussi leurs limites, notamment celle de mettre l'accent sur la quantifiabilité et de ne pas révéler immédiatement les biais potentiels des algorithmes. Les modèles de langue existants se heurtent également aux **limites inhérentes aux sources numériques, notamment celles qui sont** issues de la numérisation des textes imprimés et de leur transformation au moyen de la reconnaissance optique de caractères (OCR). Cependant, ces biais peuvent être en partie corrigés grâce aux avancées rapides dans le domaine du *deep learning*, renforcées par la création de jeux de données d'entraînement de qualité (*groundtruth* ou *gold standard*) produites manuellement par des chercheurs experts du domaine. Les campagnes d'annotations réalisées dans ce cadre illustrent le potentiel d'une intégration critique des méthodes computationnelles et conventionnelles et d'une coopération symbiotique entre sciences humaines et informatiques. Par ailleurs, bien que le « quantitatif » ait mauvaise presse dans les sciences humaines depuis le tournant culturel, la quantification a aussi ses vertus. Elle peut être un correctif salutaire aux dérives interprétatives des études culturelles aujourd'hui dominantes. Inversement, l'herméneutique est un outil indispensable pour donner sens aux résultats des analyses computationnelles.

Revenons à l'exemple de la littérature. Si les méthodes computationnelles peuvent analyser de vastes corpus textuels, elles ne peuvent évidemment pas le faire avec la sensibilité d'un lecteur touché par le style propre d'un auteur, les émotions que réveille un poème ou l'empathie éprouvée pour un personnage de roman. Le *distant reading* ne prétend pas remplacer le *close reading*. L'analyse statistique des corpus ne dilue pas la singularité des textes. La détection de régularités ne réduit pas la variété des interprétations qu'ils peuvent nourrir chez chaque lecteur. Il faut appréhender les méthodes computationnelles comme un jeu de clés ouvrant le champ de possibles, révélant des textes ou des motifs qui seraient peut-être passés inaperçus et suggérant des perspectives alternatives au lecteur trop près de son texte ou engoncé dans une grille de lecture préconçue. Ensemble, ces deux manières complémentaires d'aborder la littérature créent les conditions d'une compréhension plus holistique et plus riche de la création humaine.

Dans le domaine de la recherche historique, les historiens qui s'intéressent à des individus ou des groupes sociaux spécifiques peuvent recourir au traitement automatique du langage (TAL) pour extraire l'information biographique des sources historiques et les transformer en bases de données prosopographiques. Les liens ainsi établis entre les *data points* peuvent révéler des relations invisibles à l'œil nu ou à partir d'une lecture linéaire des sources. Les historiens peuvent alors conduire des analyses autrement impossibles, comme la reconstitution de carrières, l'analyse des réseaux ou l'identification de groupes sur la base d'attributs partagés (génération, lieu de naissance, occupation, etc.). Identifier des tendances et des régularités permet aussi de mieux situer les cas individuels voire atypiques au sein de la population étudiée et de repérer des ruptures ou des tournants fragmentant la longue durée. Dans le cadre d'une historiographie mobilisant les méthodes computationnelles, une solide connaissance du contexte historique et une lecture attentive des sources restent indispensables pour façonner les modèles, structurer les données et interpréter correctement les résultats algorithmiques, de manière à construire un récit intelligible et conforme à la réalité historique.

Dans le domaine de la sociologie ou de l'anthropologie, les chercheurs peuvent employer les méthodes computationnelles pour détecter l'émergence d'un phénomène d'opinion et en cartographier la propagation sur les réseaux sociaux. Cependant, pour évaluer efficacement l'influence des médias en ligne sur les mouvements sociaux ou politiques, les chercheurs spécialistes du terrain numérique doivent en parallèle recueillir des données empiriques sur les événements contemporains qui se produisent dans le « monde réel ». La profondeur de leur analyse dépend aussi de leur capacité à croiser leurs résultats avec ceux tirés d'autres médias plus conventionnels, tels que les journaux et la télévision.

La transformation des sciences humaines à l'ère numérique n'est pas une affaire de remplacement, mais de symbiose. Il faut commencer par reconnaître les forces et les limites de chaque approche pour les intégrer de manière à rendre compte de l'expérience humaine dans toute ses nuances et sa complexité. La voie à suivre est claire : celle d'une interdisciplinarité authentique et d'un engagement renouvelé avec les valeurs fondamentales des sciences humaines.

Conclusion : négocier le tournant computationnel

Les Humanités numériques ou computationnelles, avec leur promesse d'innovation et d'interdisciplinarité, ont laissé une empreinte indélébile sur le paysage académique. Pourtant, comme nous l'avons souligné, ces notions englobantes ne capturent qu'une fraction des changements en cours. Elles entretiennent l'illusion que les humanités numériques consistent essentiellement en une révolution technique, alors qu'il s'agit d'une mutation épistémologique bien plus profonde. L'essence de cette transformation ne réside pas dans l'adoption superficielle d'outils numériques, mais dans **l'intégration des méthodes computationnelles au cœur des différentes disciplines, dans le respect de leur spécificité.**

Chaque discipline possède son propre cadre de conventions intellectuelles, de questions et de pratiques. Les sciences computationnelles offrent des opportunités uniques de remodeler ces traditions disciplinaires. Allant du *distant reading* appliqué aux vastes corpus littéraires aux analyses de réseaux sociaux ou historiques, l'éventail des méthodes disponibles ouvre des horizons de recherche quasi infinis. Pourtant, leur véritable potentiel ne peut être atteint que si elles sont entrelacées avec les méthodes établies par chaque discipline, de manière à créer une relation symbiotique susceptible d'amplifier leurs forces respectives.

Nous soutenons que l'élément clé pour faciliter cette transition est de s'éloigner des concepts globaux (DH, CH, CSS, etc.) et de **mettre l'accent sur les méthodes** concrètes disponibles aux chercheurs en sciences humaines. L'éventail de ces méthodes est vaste. S'il n'est pas nécessaire de toutes les utiliser en même temps, ni avec le même degré d'implication, il est essentiel de pouvoir discerner quelles sont les plus appropriées à leurs questionnements, à leur approche ou à la nature de leurs sources. Les chercheurs doivent partir de leurs besoins spécifiques pour identifier les méthodes les plus susceptibles de correspondre à leur quête intellectuelle. L'historien, le littéraire, le sociologue ne devraient pas se demander ce que peuvent les « humanités numériques », mais plutôt s'efforcer de comprendre ce que les différentes méthodes computationnelles peuvent individuellement et conjointement apporter à leurs recherches.

Les chercheurs doivent embrasser le tournant computationnel non pas comme une révolution, mais comme **une extension critique des valeurs fondamentales des sciences humaines**. Les sciences humaines ont toujours prôné la profondeur, la nuance et le sens critique. Lorsqu'elles sont utilisées avec discernement, les méthodes computationnelles viennent amplifier ces valeurs, offrant à la fois précision, rigueur, changements d'échelle et de perspectives. Il ne s'agit pas simplement d'adopter de nouveaux outils mais d'acquérir une compréhension intime du paysage computationnel. Si l'historien (et l'humaniste en général) est appelé à devenir programmeur, ce n'est pas pour créer lui-même des algorithmes (d'autres le font mieux que lui) mais pour s'engager de manière critique avec les algorithmes existants afin de comprendre leurs biais et s'assurer que leur application est en accord avec les fondements éthiques et épistémologiques des sciences humaines. Pour ce faire, les chercheurs en SHS doivent travailler en étroite collaboration avec des spécialistes en informatique, pas nécessairement au quotidien, mais à différents stades de leurs recherches. Cet engagement interdisciplinaire implique de dépasser les pratiques actuelles d'**autoformation pour mettre en place des formations en sciences computationnelles adaptées et intégrées aux parcours conventionnels des humanistes au sein même des institutions académiques.**

L'un des principaux obstacles est en effet **l'absence totale de formation aux méthodes computationnelles pour les chercheurs en sciences humaines** dans la plupart des universités,

alors même que les institutions académiques devraient avoir pour rôle d'encourager et de faciliter cette transition. Un deuxième défi connexe est le **manque de reconnaissance** que reçoivent généralement les compétences computationnelles dans les procédures de recrutement ou de promotion académiques. De même, l'usage raisonné de ces méthodes est rarement pris en compte lors du processus d'évaluation par les pairs dans les publications académiques conventionnelles. À cet égard, le [Journal of Digital History](#) constitue une exception notable dans le domaine de la recherche historique, qui mérite d'être saluée.

Pour conclure, nous appelons au développement d'un nouveau mode de pensée computationnelle dans les sciences humaines, qui repose sur **trois qualités essentielles** : la capacité à traduire les questions des sciences humaines en tâches exécutables par l'ordinateur ; l'aptitude à interpréter les résultats computationnels dans une perspective humaniste et critique ; et enfin, la capacité à intégrer ces nouveaux résultats aux connaissances antérieures, de manière à construire un récit cohérent susceptible non seulement de faire avancer l'état des connaissances dans le domaine concerné, mais aussi de toucher une plus large communauté de lecteurs au-delà de ce domaine spécifique. Ce sont trois compétences fondamentales que les chercheurs en sciences humaines devront acquérir dans un avenir proche s'ils entendent conserver leur pertinence à l'ère de l'intelligence artificielle.

Le dépassement des humanités numériques vers les sciences computationnelles n'est donc pas un renoncement aux sciences humaines, mais plutôt un approfondissement de leur essence. C'est en exploitant pleinement mais avec discernement la puissance des algorithmes que l'on peut peut-être, pour la première fois dans l'histoire des sciences, envisager de saisir l'expérience humaine dans toute sa complexité.

Références

- Abanoz, Enes, dir. *Opportunities and Challenges for Computational Social Science Methods*. Advances in Human and Social Aspects of Technology Book Series. Hershey, PA: Information Science Reference, 2022.
- Alvarez, R. Michael, and Seo-young Silvia Kim. *Computational Social Science*. Oxford Bibliographies. New York: Oxford University Press, 2019.
- Alvarez, R. Michael, éd. *Computational Social Science: Discovery and Prediction*. Analytical Methods for Social Research. New York: Cambridge University Press, 2018.
- Armand, Cécile. "Bonding minds, bridging nations: Sino-American alumni networks in the Era of Exclusion (1882-1936)", in Henriot (dir.), *Modern China in Flux: Networks, Mobility, and Transformation*, Berlin, De Gruyter (scheduled for publication in 2024).
- Armand, Cécile. "Mapping the Transnational Public Sphere in Modern China: A Bilingual Topic Modeling of the Republican Press (1919-1949)." *Journal of Digital History*, Fall 2023.
- Bay, Jessica. *Matthew K. Gold, Editor. Debates in the Digital Humanities*, University of Minnesota Press, 2012, Vol. 4, 2014.
- Bertoni, Eleonora, Matteo Fontana, Lorenzo Gabrielli, Serena Signorelli, and Michele Vespe, dir. *Handbook of Computational Social Science for Policy*. Cham: Springer International Publishing, 2023.
- Blouin Baptiste, Cécile Armand, Christian Henriot. HistText: An Application for leveraging large-scale historical textbases. 2023. [⟨halshs-04178820⟩](#)
- Bod, Rens, Jaap Maat, and Thijs Weststeijn, dir. *The Making of the Humanities. Volume III, The Modern Humanities*. Amsterdam: Amsterdam University Press, 2014.

- Bode, Katherine. « The Equivalence of “Close” and “Distant” Reading; or, Toward a New Object for Data-Rich Literary History ». *Modern Language Quarterly* 78, n° 1 (1 mars 2017): 77-106.
- Bodenhamer, David J., John Corrigan, and Trevor M. Harris, dir. *The Spatial Humanities: GIS and the Future of Humanities Scholarship*. Spatial Humanities (Indiana University Press). Bloomington: Indiana University Press, 2010.
- Burdick, Anne, Johanna Drucker, Peter Lunenfeld, Todd Samuel Presner, and Jeffrey T. Schnapp. *Digital-Humanities*. Cambridge, Mass.: The MIT Press, 2016.
- Chambliss, Julian C., and Ellen Moll. *Making Sense of Digital Humanities: Transformations and Interventions in Technocultures*. East Lansing: MSU Libraries, 2022.
- Chen, Shu-Heng, dir. *Big Data in Computational Social Science and Humanities*. Computational Social Sciences. Cham: Springer International Publishing, 2018.
- Cohen, Daniel J, and Roy Rosenzweig. *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*. Philadelphia: University of Pennsylvania Press, 2006.
- Engel, Uwe, Anabel Quan-Haase, Sunny Xun Liu, and Lars Lyberg, dir. *Handbook of Computational Social Science. Volume 1, Theory, Case Studies and Ethics*. European Association of Methodology Series. London: Routledge, 2022.
- Engel, Uwe, Anabel Quan-Haase, Sunny Xun Liu, and Lars Lyberg, dir. *Handbook of Computational Social Science. Volume 2, Data Science, Statistical Modelling, and Machine Learning Methods*. European Association of Methodology Series. London: Routledge, 2022.
- Fickers, Andreas, and Juliane Tatarinov, dir. *Digital History and Hermeneutics: Between Theory and Practice*. De Gruyter, 2022.
- Fiormonte, Domenico, Francesca Tomasi, and Teresa Numerico. *The Digital Humanist: A Critical Inquiry*. Translated by Christopher Ferguson and Desmond Schmidt. Baltimore, Maryland: Project Muse, 2020.
- Fitzpatrick, Kathleen. *Planned Obsolescence Publishing, Technology, and the Future of the Academy*. New York: New York University Press, 2012.
- Foster, Ian, Rayid Ghani, Ron S. Jarmin, Frauke Kreuter, and Julia Lane. *Big Data and Social Science: Data Science Methods and Tools for Research and Practice*. Sous la direction de Ian Foster, Rayid Ghani, Ron S. Jarmin, Frauke Kreuter, and Julia Lane. 2nd ed. Chapman and Hall/CRC, 2020.
- Fu Haiyan, Xu Jian, dir. *Dashuju yu zhongguo lishi yanjiu*. Beijing : Shehui kexue wenxian chubanshe , vol.1, 2017, vol. 2, 2020, Vol. 3, 2021, Vol. 4, 2023
- Gilbert, G. Nigel. *Computational Social Science*. Sage Benchmarks in Social Research Methods. London: SAGE, 2010.
- Gillet, Seth, Florence Hengchen, and Simon Max deWilde van Hooland. *Introduction Aux Humanités Numériques Méthodes et Pratiques*. Louvain-la-Neuve : De Boeck Supérieur, 2016.
- Goldstone, Andrew, et Ted Underwood. « The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us ». *New Literary History* 45 (1 juin 2014): 359-84.
- Gregory, Ian N., and A. Geddes, dir. *Toward Spatial Humanities: Historical GIS and Spatial History*. Spatial Humanities (Indiana University Press). Bloomington: Indiana University Press, 2014.
- Gregory, Ian, Donald A. DeBats, and Donald Lafreniere, eds. *The Routledge Companion to Spatial History*. Milton Park, London: Routledge, 2018.

- Grimmer, Justin, Margaret E. Roberts, and Brandon M. Stewart. *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton: Princeton University Press, 2022.
- Henriot, Christian. "Eminent Chinese of the Shenbao (1872-1891). A digital investigation of news reporting and newspaper-making in late imperial China." *Journal of Digital History*, Fall 2023.
- Humanités numériques : de nouveaux récits en histoire de l'art ?* Histoire de l'art. Paris : Association des professeurs d'archéologie et d'histoire de l'art des universités (APAHU), 2021.
- Jockers, Matthew Lee. *Macroanalysis: Digital Methods and Literary History*. Topics in the Digital Humanities. Urbana: University of Illinois Press, 2017.
- Jones, Steven E., *The Emergence of the Digital Humanities*. London: Taylor and Francis, 2020.
- Kee, Kevin, dir. *Pastplay: Teaching and Learning History with Technology*. Ann Arbor: University of Michigan Press, 2014.
- Kertész, János, Rosario N. Mantegna, and Salvo Micciché, dir. *Computational Social Science and Complex Systems*. Proceedings of the International School of Physics "Enrico Fermi." Amsterdam: IOS Press, 2019.
- Kim, Dorothy, and Jesse Stommel, dir. *Disrupting the Digital Humanities*. Goleta, CA: Punctum Books, 2018.
- Lazer, David, and Stefan Wojcik. *Political Networks and Computational Social Science*. Edited by Jennifer Nicoll Victor, Alexander H. Montgomery, and Mark Lubell. Vol. 1. Oxford: Oxford University Press, 2017.
- Liu Zhiwei, Wang Lei, dir. *Shuzi renwen yu xin wenke fazhan*. Beijing : Shehui kexue wenxian chubansh , 2023
- Manifesto of Computational Social Science*. EDP Sciences, 2012.
- McLevey, John. *Doing Computational Social Science: A Practical Introduction*. Los Angeles: SAGE, 2022.
- Moretti, Franco. *Graphs, Maps, Trees: Abstract Models for a Literary History*. ACLS Humanities E-Book. London: Verso, 2007.
- Muhanna, Elias, ed. *The Digital Humanities and Islamic & Middle East Studies*. Berlin: De Gruyter, 2016.
- Nelimarkka, Matti. *Computational Thinking and Social Science: Combining Programming, Methodologies and Fundamental Concepts*. Los Angeles: SAGE, 2023.
- Payne, Diane, Johan A. Elkink, Nial Friel, Thomas U. Grund, Tamara Hochstrasser, Pablo Lucas, and Adrian Ottewill. *Social Simulation for a Digital Society: Applications and Innovations in Computational Social Science*. Springer Proceedings in Complexity. Cham: Springer, 2020.
- Ramsay, Stephen. *Reading Machines: Toward an Algorithmic Criticism*. Topics in the Digital Humanities. Urbana: University of Illinois Press, 2017.
- Ridolfo, Jim. *Digital Samaritans: Rhetorical Delivery and Engagement in the Digital Humanities*. Digital Rhetoric Collaborative (Ann Arbor, Mich.). Ann Arbor: University of Michigan Press, 2015.
- Rudas, Tamás, and Gábor Péli, dir. *Pathways Between Social Science and Computational Social Science: Theories, Methods, and Interpretations*. Computational Social Sciences. Cham: Springer International Publishing, 2021.
- Salganik, Matthew J. *Bit by Bit: Social Research in the Digital Age*. Princeton: Princeton University Press, 2019.

- Schiuma, Giovanni, and Daniela Carlucci, dir. *Big Data in the Arts and Humanities: Theory and Practice*. Data Analytics Applications. Boca Raton, FL: CRC Press, Taylor & Francis Group, 2018.
- Schreibman, Susan, Raymond George Siemens, et John Unsworth. *A Companion to Digital Humanities*. Blackwell Companions to Literature and Culture. Malden, MA: Blackwell Pub., 2004.
- Schreibman, Susan, Raymond George Siemens, et John Unsworth. *A New Companion to Digital Humanities*. Blackwell Companions to Literature and Culture. Malden, MA: Chichester: Wiley/Blackwell, 2016.
- Shah, Dhavan V., Joseph N. Cappella, and W. Russell Neuman, dir. *Toward Computational Social Science: Big Data in Digital Environments*. Annals of the American Academy of Political and Social Science. Los Angeles: SAGE, 2015.
- Shixue yuekan, *Dashuju shidai de shiliao yu shixue*. Beijing: Renmin chubanshe, 2017.
- Shu Jian, ed. *Dashuju shidai de lishi yanjiu*. Shanghai : Shanghai yiwen chubanshe, 2017.
- Shuwei renwen yu jindai zhongguo zhishifenzi*, Zhengzhi daxue shuwei shiliao yanjiu congkan. Taipei: Guoli zhengzhi daxue tushuguan, 2020.
- Siemens, Raymond George, and Susan Schreibman. *A Companion to Digital Literary Studies*. Blackwell Companions to Literature and Culture. Chichester: Wiley-Blackwell, 2013.
- Sophie Anquetil, Carine Duteil-Mougel, Vivien Lloveria. *Le Sens des données. Le Statut du corpus et herméneutique à l'aune des Humanités Numériques*. Paris: Editions L'Harmattan, 2019.
- Svensson, Patrik, and David Theo Goldberg, dir. *Between Humanities and the Digital*. Cambridge, Mass.: MIT Press, 2015.
- The Shape of Data in the Digital Humanities Modeling Texts and Text-Based Resources*. Digital Research in the Arts and Humanities. London: Routledge, 2020.
- Travis, Jennifer, and Jessica DeSpain, dir. *Teaching with Digital Humanities: Tools and Methods for Nineteenth-Century American Literature*. Topics in the Digital Humanities. Urbana: University of Illinois Press, 2018.
- Underwood, Ted. *Distant Horizons: Digital Evidence and Literary Change*. Chicago, IL: University of Chicago Press, 2019.
- Vinck, Dominique. *Humanités Numériques : La culture face aux nouvelles technologies*. Paris: Le cavalier bleu, 2016.
- Whalen, Ryan, ed. *Computational Legal Studies: The Promise and Challenge of Data-Driven Research*. Elgar Studies in Legal Research Methods. Cheltenham: Edward Elgar Publishing Limited, 2020.
- Xiang Jie, ed. *Cong baocun dao chuangzao: kaiqi shuwei renwen yanjiu* (From preservation to knowledge creation: the way to digital humanities). Shuwei renwen yanjiu congshu (Taipei: Guoli taiwan daxue chuban zhongxin, 2011 (reed. 2016).
- Xiang Jie. *Shuwei renwen yanjiu yu jiyi* (Digital humanities and craft: technological change). Series on digital humanities. Taipei : Guoli taiwan daxue chuban zhongxin, 2014.
- Yaovi Akakpo. *Humanités Numériques et Éducation en Afrique. Innovations sociales en Afrique*. Paris : L'Harmattan, 2021.
- Ye, Xinyue, and Hui Lin, dir. *Spatial Synthesis: Computational Social Science and Humanities*. Human Dynamics in Smart Cities. Cham: Springer International Publishing, 2020.