



Humanités numériques et modélisation scientifique

Jean-Guy Meunier

DANS **QUESTIONS DE COMMUNICATION 2017/1 n° 31**, PAGES 19 À 48

ÉDITIONS **ÉDITIONS DE L'UNIVERSITÉ DE LORRAINE**

ISSN 1633-5961

DOI 10.4000/questionsdecommunication.11040

Date de mise en ligne : 04/09/2017

Article disponible en ligne à l'adresse

<https://shs.cairn.info/revue-questions-de-communication-2017-1-page-19?lang=fr>



Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...
Scannez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.



Distribution électronique Cairn.info pour Éditions de l'Université de Lorraine.

Vous avez l'autorisation de reproduire cet article dans les limites des conditions d'utilisation de Cairn.info ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Détails et conditions sur cairn.info/copyright.

Sauf dispositions légales contraires, les usages numériques à des fins pédagogiques des présentes ressources sont soumises à l'autorisation de l'Éditeur ou, le cas échéant, de l'organisme de gestion collective habilité à cet effet. Il en est ainsi notamment en France avec le CFC qui est l'organisme agréé en la matière.

Humanités numériques et modélisation scientifique

Digital Humanities and Scientific Modelling

Jean-Guy Meunier



Édition électronique

URL : [http://
questionsdecommunication.revues.org/11040](http://questionsdecommunication.revues.org/11040)
DOI : 10.4000/
questionsdecommunication.11040
ISSN : 2259-8901

Éditeur

Presses universitaires de Lorraine

Édition imprimée

Date de publication : 1 septembre 2017
Pagination : 19-48
ISBN : 9782814303256
ISSN : 1633-5961

Distribution électronique Cairn



CHERCHER, REPÉRER, AVANCER.

Référence électronique

Jean-Guy Meunier, « Humanités numériques et modélisation scientifique », *Questions de communication* [En ligne], 31 | 2017, mis en ligne le 01 septembre 2019, consulté le 04 septembre 2017. URL : <http://questionsdecommunication.revues.org/11040> ; DOI : 10.4000/questionsdecommunication.11040

Tous droits réservés

JEAN-GUY MEUNIER
Département de philosophie
Université du Québec à Montréal
CA-H3C 3P8
meunier.jean-guy@uqam.ca

HUMANITÉS NUMÉRIQUES ET MODÉLISATION SCIENTIFIQUE

Résumé. — Les projets dans les humanités numériques sont-ils de nature scientifique ? Une vision dite pragmatique et cognitive de la science montre que les théories scientifiques sont avant tout des modèles de différents types qui, ensemble, contribuent à l'avancement du savoir. Implicitement ou explicitement, les humanités numériques utilisent des modèles formels, matériels, conceptuels qui leur permettent de se positionner comme démarche scientifique, mais qui, aussi, limitent et contraignent leur méthode et leur champ d'application dans le domaine des humanités. En ce sens, les humanités numériques créent une passerelle originale entre les sciences et l'herméneutique.

Mots clés. — humanités numériques, modélisation, philosophie des sciences, épistémologie

Dans les années 2000, lié à l'émergence des technologies informatiques, est apparu un programme de recherche original et novateur habilement appelé les humanités numériques calquant l'anglais *Digital Humanities*. Ce programme reçut plusieurs définitions. En voici trois :

« Par *digital humanities*, nous désignons les recherches qui intègrent les technologies de l'information comme partie centrale de leur méthodologie pour créer et/ou traiter des données. [...] L'emploi du terme renvoie au constat de l'importance croissante que les outils et ressources numériques ont maintenant pour les sciences sociales » (nous traduisons)¹.

« Les humanités numériques comprennent l'étude de ce qui se passe à l'intersection des outils informatiques avec des artefacts culturels de toute sorte »².

« Les humanités numériques recouvrent un ensemble de pratiques de recherche à l'intersection des technologies numériques et des différentes disciplines des sciences humaines » (Dacos, Mounier, 2014 : 7).

Les deux termes *humanities* et *humanités* ne désignent pas les mêmes objets de recherche. En anglais, *humanities* désigne traditionnellement les sciences liées à la culture telles la sociologie, l'anthropologie culturelle, les sciences du langage, le droit, les pratiques artistiques (musique, théâtre, peinture, danse), la littérature et même le journalisme. Lorsque la signification est un peu plus précise, les *humanities* incluent plusieurs disciplines des sciences humaines, par exemple la sociologie, la philosophie, les communications médiatiques. Dans le monde francophone, italien, espagnol, les *humanités* désignent généralement plus une tradition intellectuelle et une éthique *humaniste* d'inspiration érasimienne qu'un champ disciplinaire universitaire. Mais lorsque cette tradition s'exprime dans des formes sémiotiques, la recherche s'intéresse à des artefacts relevant des arts, des lettres, des médias et même de certaines sciences (médicale, juridique, cognitive, politique, sociologique, etc.).

Cependant, on accepte de plus en plus le recouvrement des deux extensions des termes *humanities* et *humanités*, surtout si l'on accole les qualificatifs *digital* au premier et *numérique* au second. Et, dans certains cas, on inclut en plus les *digital studies*, les systèmes d'information et de communication et toute la question de l'instrumentation numérique de la connaissance. Pour les fins de notre analyse, nous utilisons ce néologisme, *humanités numériques* (ci-après HN), pour désigner de manière très générale un territoire de recherche qui, malgré de multiples variantes, met en relation deux entités radicalement différentes³. D'un côté, on trouve des disciplines des sciences humaines et sociales dites des *humanities* ou *humanités* qui

¹ Dans le présent article, sauf mention contraire, nous traduisons. « By *digital humanities*, we mean research that uses information technology as a central part of its methodology, for creating and/or processing data...The use of the term *digital Humanities*] reflects a growing sense of the importance that digital tools and resources now have for humanities subjects ». Accès : <http://digital.humanities.ox.ac.uk/Support/whatarethedh.aspx>. Consulté le 26/05/2017.

² « The *digital humanities* comprise the study of what happens at the intersection of computing tools with cultural artefacts of all kinds ». Accès : <http://www.digitalhumanities.org/dhq/about/about.html>. Consulté le 26/05/2017.

³ Sur cette question, voir J. Bonaccorsi, V. Carayol et J.-C. Domenget (2016). On y présente des réflexions sur les nouveaux territoires disciplinaires que définit l'instrumentation de la connaissance par le numérique.

s'intéressent à des objets complexes de types sémiotiques, par exemple, les lettres, les langues et les communications, la culture, les représentations sociales, etc. ; de l'autre, on trouve des technologies dites numériques de traitement d'information.

Malgré la résistance qu'a soulevée la rencontre de ces deux qualificatifs, *humanités* et *numériques*, et le haut de degré de généralité de ces définitions, il faut bien voir qu'il s'agit d'un programme de recherche en sémiotique computationnelle (Meunier, à paraître), c'est-à-dire d'une approche qui modélise de manière computationnelle les systèmes de signes et de symboles. En effet, les HN manipulent des symboles ou des représentations sémiotiques, c'est-à-dire – pour reprendre le vocabulaire classique peircien – des signes, mais cette manipulation est effectuée à l'aide de la technologie dite numérique. Ces symboles sont signifiants et de types variés : textes, sons, images, comportements, etc. Et le traitement est réalisé par des ordinateurs.

Dans de tout autres vocabulaire et paradigme, il s'avère que l'intelligence artificielle (IA) s'est définie de manière très similaire : en effet, comme l'a répété Allen Newell (1983 : 196), « les scientifiques de l'IA voient les ordinateurs comme des machines qui manipulent des symboles »⁴. Ces derniers ayant aussi une signification, c'est-à-dire qu'ils désignent quelque chose : « Le concept le plus fondamental pour un symbole est ce qui lui confère le caractère symbolique, à savoir ce qui lui permet de représenter [*stand for*] une certaine entité. Nous appelons cela des concepts » (Newell, 1980 : 156)⁵.

Ainsi, d'un point de vue épistémologique, peut-on voir les humanités numériques comme des types spécifiques de projets d'intelligence artificielle, mais où les « symboles » sont ceux qui, cette fois, désignent des réalités qui intéressent diverses sciences humaines et sociales (SHS) et que l'on veut traiter de manière computationnelle.

Cependant, une telle vision des HN rencontre deux positions antagonistes : d'un côté, celle qui, venant des SHS, consiste à voir l'ordinateur comme un simple outil-gadget qui ne peut effectuer que des traitements superficiels en raison de la complexité des objets sémiotiques étudiés ; de l'autre, celle qui, venant de l'informatique, empêche de croire à la pertinence du traitement informatique pour des objets difficiles à circonscrire et à saisir. Au contraire, certains promoteurs des humanités numériques voient dans les HN l'ouverture d'un programme de recherche original renouvelant de manière radicale la pratique et les méthodes mêmes des SHS et des humanités.

Notre objectif est de sortir de ces positions critiques ou prophétiques pour tenter de comprendre, selon une approche spécifique de la philosophie des sciences dite pragmatique, la nature de la relation entre les humanités et le « numérique ». Nous explorerons l'hypothèse selon laquelle les humanités numériques sont des démarches de connaissances qui, à l'instar des pratiques scientifiques, construisent des théories et, d'une part, visent à décrire, expliquer et comprendre la complexité de certains objets

⁴ « *Scientist in AI saw computers as machines that manipulated symbols* ».

⁵ « *The most fundamental concept for a symbol is that which gives symbols their symbolic characters, i.e. which lets them stand for some entity. We call this concepts* ».

sémiotiques et, d'autre part, sont assistées par des traitements dits « numériques », applicables à ces objets sémiotiques. Nous espérons montrer que les HN permettent de penser une rencontre originale entre une démarche de type scientifique d'influence pragmatique et une démarche scientifique d'influence herméneutique liée à l'interprétation requise par la sémiotité de ses objets de recherche.

La théorie scientifique

À première vue, il semble difficile de placer les HN dans le cadre d'une recherche de type scientifique car on les situerait intuitivement dans l'horizon des sciences herméneutiques. Mais cette difficulté est tributaire d'une vision spécifique de la science qui est pensée comme un ensemble d'énoncés formels dont la vérité est validée par l'expérimentation ; ce qui constitue l'essentiel de sa valeur aléthique et explicative.

Cette vision restrictive de la science a été remise en question par plusieurs scientifiques et philosophes épistémologues, non parce qu'elle est fautive, mais parce qu'elle est normative plutôt que pragmatique. À ce titre, elle est souvent réductrice et privilégie la démonstration par l'expérimentation. Mais, surtout, elle ne correspond pas aux pratiques réelles de plusieurs disciplines scientifiques contemporaines.

L'une des thèses importantes de la philosophie classique des sciences est que la science est un discours qui construit une *théorie* à propos de ses objets. Ce n'est pas tant cette définition de la science comme discours qui fait problème que la compréhension du concept même de *théorie*. De fait, c'est sur la nature d'une telle théorie qu'il y a divergence de points de vue. Divergence qui entraîne une révision de la compréhension de la nature de la science.

Rappelons brièvement les différentes conceptions que la philosophie des sciences propose sur la nature d'une *théorie* dite scientifique. Une première, appelée *syntaxique*, vient du positivisme logique et définit une théorie comme un ensemble d'énoncés formés par des règles strictes et dont certains sont soit des théorèmes, soit des axiomes ou des postulats :

« La théorie scientifique [...] est entendue comme un ensemble constitué de formules théoriques syntaxiquement formées (axiomes, théorèmes et lois) et de leur interprétation sous forme d'énoncés de correspondance » (Winther, 2015 : en ligne)⁶.

Dite *sémantique*, une deuxième vision reprend la même position, mais précise la nature des énoncés. Pour Bas C. van Fraassen (1970 : 327), une théorie doit être « une structure mathématique ». Autrement dit, ce n'est pas qu'un ensemble d'énoncés syntaxiquement bien formés, mais aussi une sémantique se rapportant avant tout à des structures mathématiques pertinentes pour un domaine de recherche.

⁶ « *Scientific theory is [...] taken to be a syntactically formulated set of theoretical sentences (axioms, theorems, and laws) together with their interpretation via correspondence sentences* ».

Ces deux visions ont été abondamment critiquées. Bien que ces types de théories soient techniquement acceptables, elles demeurent rares et peu applicables dans la majorité des sciences (Papineau, 2010 ; Leplin, 1984). Des études analytiques ont montré que de multiples pratiques scientifiques ne se réduisent pas à des théories comme ensembles d'énoncés formels et que le type d'explication qu'elles fournissent ne se présente pas uniquement sous forme nomologique et déductive. C'est souvent ces dernières que les acteurs des humanités/*humanities* ont en tête lorsqu'on veut critiquer la scientificité de leur démarche.

Une troisième vision dite *pragmatique* est apparue, influencée par la longue tradition du pragmatisme américain : William James, John Dewey, Wilfrid Sellars, Paul Feyerabend, Thomas Kuhn, Ian Hacking, etc. Elle n'est pas sans lien avec des courants européens de sciences de la culture et de la sociologie des connaissances (Max Weber, Karl Mannheim, Michel Foucault, Thomas Luckman, François Rastier, etc.). Et, à la suite des relectures entre autres auteurs de Pierre Duhem (1906) et Suzanne Bachelard (1979), elle fut revisitée par Hans-Jörg Rheinberger (1997), Ronald N. Giere (1999), Nancy Cartwright (1983), Mary S. Morgan et Margaret Morrison (1999). Cette vision pragmatique a été construite à partir d'études de plusieurs pratiques concrètes, tant en sciences de la nature qu'en sciences humaines et même en ingénierie. Une théorie y est vue comme un moment, un état de stabilisation d'un processus cognitif complexe souvent itératif, qui s'exprime sous diverses formes sémiotiques et discursives, formelles ou non formelles, ou même iconiques (les graphes, les plans, les simulations, etc.). Ce processus cognitif réunit divers points de vue, perspectives et stratégies d'enquête, démonstrations, langages, etc. Dans la vision pragmatique, chacun de ses points de vue, angles, perspectives, intérêts est regroupé sous le terme englobant *modèle*.

La modélisation en science

Cette notion de modèle n'est pas nouvelle. Dans ses premières utilisations, elle fut surtout associée aux sciences naturelles classiques et à la logique mathématique. Mais, comme l'ont montré Michel Armatte et Amy Dahan Dalmedico (2004), elle s'est fortement adaptée aux multiples changements survenus dans les pratiques de toutes les sciences. De plus, l'introduction de l'ordinateur en a accéléré l'adoption puisque, dès lors, toutes les sciences ont rapidement été confrontées à la nécessité de traiter de nouveaux objets, des points de vue différents, des expertises hybrides, des langages formels diversifiés. Ces changements ont immédiatement rendu cette notion de modèle des plus heuristiques et son concept encore plus important pour unifier des pratiques scientifiques utilisant des approches hybrides et diversifiées. Ainsi les modèles seront-ils vus comme des pierres d'assise de l'explication scientifique.

Des philosophes de la science (Cartwright, 1983 ; Le Moigne, 1987 ; Frigg, Hartmann, 2006 ; Knuuttila, 2011) donnent au concept de modèle un rôle épistémique essentiel dans toutes les sciences, qu'elles soient naturelles, sociales ou humaines : « Fournir une explication pour un phénomène c'est trouver un modèle qui s'ajuste au cadre de

base d'une théorie et qui permet des analogies avec des lois phénoménologiques peu soignées et compliquées mais qui sont vraies pour ce modèle » (Cartwright, 1983 : 22)⁷. Une théorie scientifique n'est plus simplement un ensemble d'énoncés, mais une collection de plusieurs types de modèles contenant divers types d'énoncés et qui tous et collectivement participent à la connaissance d'une réalité étudiée. On rejoint donc l'intuition souvent citée de Paul Valéry : « Nous ne raisonnons que par modèles ».

Le rôle cognitif que les modèles jouent dans une théorie scientifique est original. Les modèles assistent la théorie dans la multiplicité et la diversité de leurs tâches et leurs objets de recherche, par exemple : la conception de la problématique, la formulation des hypothèses, l'intégration aux recherches existantes, la définition de la méthode, l'introduction d'une formalisation, l'illustration par la simulation, la construction de prototypes, le design de l'expérimentation, l'interprétation et la dissémination des résultats, etc. Mary S. Morgan et Margaret Morrison (1999 : 15) résument ce rôle par une formule : un modèle sert à « assembler des petits bouts qui viennent de sources disparates »⁸. Mais, elles précisent que les modèles sont des médiateurs « entre les théories et le monde » (« *mediate between our theories and the world* », *ibid.* : 17). Pour résumer, cette notion est devenue une thématique importante de la philosophie des sciences (Burch, 1999 ; Frigg, Hartmann, 2006).

Cependant, une définition technique de ce qu'est un modèle demeure difficile à identifier : souvent, on se contente de les définir de manière extensionnelle, c'est-à-dire en donnant des exemples ou des cas particuliers. Mais on trouve aussi des définitions intensionnelles où apparaissent les sens divers qu'on donne au concept : « En science, le terme *modèle* a plusieurs sens et ces différents types de modèles servent des fins scientifiques distinctes » (Craver, 2006 : 355 ; Giere, 1999 : 1)⁹. Pour Tudor M. Băetu (2013 : 2), ils sont des expressions des perspectives cognitives diverses que l'on peut avoir sur l'objet d'étude :

« Les modèles ancrent les divers éléments de la mosaïque de la connaissance, d'un côté, à la description d'un phénomène et, de l'autre, aux méthodes et outils – expérimentaux ou théoriques – utilisés pour obtenir chacun de ces éléments »¹⁰.

Les modèles ne sont donc pas des théories, mais des outils visant à construire des théories qui, elles, exigent un engagement ontologique et aléthique. Hans-Jörg Rheinberger (1997) écrit que les modèles ne servent pas tant à *explorer ce que l'on sait, mais plutôt ce que l'on ignore*. Autrement dit, ce sont des artefacts épistémiques

⁷ « To explain a phenomenon is to find a model that fits it into the basic framework of the theory and that thus allows us to derive analogues for the messy and complicated phenomenological laws which are true of it ».

⁸ « Fitting together a set of bits which come from disparate sources ».

⁹ « There are many senses of the term "model" in science, and these different kinds of model serve distinct scientific ends ».

¹⁰ « Models anchor the diverse pieces of the mosaic of knowledge to a description of a phenomenon, on the one side, and to the methods and tools, experimental or theoretical, used to obtain each piece of the mosaic, on the other ».

qui participent à la description, l'explication et la compréhension des objets de recherche. Ils s'expriment par diverses formes sémiotiques qui tantôt sont des langues formelles, tantôt des langues naturelles, tantôt des langues iconiques.

Quelques fois, ils relèvent d'une intuition ou d'une pratique disciplinaire très assimilée, mais peu thématifiée. D'autres fois, ils participent d'une exploration heuristique prenant la forme de protocoles, d'expérimentations, de démonstrations rigoureuses, d'illustrations, de critiques, de révisions, de prototypes et même de technologies, etc. Une telle vision permet de voir que, malgré des différences radicales, les sciences dites naturelles et celles herméneutiques ou interprétatives comportent des similarités, en ce que ces deux types de science créent de multiples modèles dans la construction de leur théorie. Et ces derniers constituent des formes particulières ou des moments dans l'interprétation de la réalité.

Comme de nombreux épistémologues (Morgan, Morrison 1999 ; Leonelli, 2007 ; Knuuttila, 2011) ayant étudié des pratiques scientifiques l'ont souligné, cette multiplicité de modèles existe dans plusieurs types de sciences telles les sciences cognitives. En effet, en raison de la multiplicité et de la diversité des tâches, et même des objets de recherche, l'utilisation de modèles devient un support heuristique essentiel. Par exemple, ces sciences étudient des opérations cognitives complexes et classiquement caractérisées par de la sémiotité ou de l'intentionnalité. Ainsi s'intéressent-elles à une multiplicité d'objets tels l'imagination, la catégorisation, le raisonnement, l'émotion, la décision, etc. En raison de la diversité, complexité et multiplicité des points de vue, des perspectives qu'on peut adopter sur ces objets, mais aussi en raison de l'utilisation de l'ordinateur dans la recherche, des auteurs parmi lesquels Daniel C. Dennett (1978), Zenon W. Pylyshyn (1984) et David Marr (1982) proposèrent que l'on considère ces sciences comme un travail de reconstruction, c'est-à-dire comme une sorte de rétro-ingénierie reposant sur trois « stances » ou postures, angles, perspectives, points de vue, prises de position, etc. sur leurs objets. Ils ont qualifié ces postures *intentionnelles*, *fonctionnelles* et *physiques*.

Inspiré par cette catégorisation, et en ajustant leur proposition, nous comprenons ces postures comme des types de *modèles* ou encore comme des classes de modèles. Dans notre taxonomie, comme nous le montrerons plus loin, ces *stances* ou postures sont des sous-modèles de chacune de trois classes de modèles que nous reconnaissons être *conceptuelles*, *formelles*, *matérielles*.

L'un des pionniers des HN, Willard McCarty (2008), reprend cette vision pragmatique et soutient que, en raison de l'utilisation de l'ordinateur, le concept de modèle joue un rôle essentiel dans la pratique des HN : « Les ordinateurs sont essentiellement des machines modélisatrices, non des juke-box de connaissances » (McCarty, 2008 : en ligne)¹¹. Cette notion apparaît aussi des plus heuristiques pour comprendre les démarches effectuées dans les HN. Et, dans le présent article, nous explorerons cette thèse épistémologique sur la théorie comme structure

¹¹ « *Computers are essentially modeling machines, not knowledge jukeboxes* ».

symbolique qui construit des modèles et nous tenterons de préciser davantage comment cette notion de modèle est pertinente pour comprendre les HN dans une perspective de pratique scientifique. Ce faisant, nous considérons les HN selon une approche relevant de la philosophie pragmatique des sciences. À leur manière, elles produisent des théories et modèles qui participent à l'explication et à la compréhension de leurs objets propres.

Une telle vision semble heuristique en ce qu'elle se glisse entre, d'une part, des positions formalistes extrêmes qui ne reconnaissent dans les HN qu'une démarche intuitive dénuée de rigueur et, d'autre part, les positions herméneutiques aussi extrêmes qui ne voient en elles que des démarches fondées sur des outils non pertinents pour assister l'interprétation. Étudions maintenant la nature et le fonctionnement de ces modèles, plus spécifiquement ceux des modèles formels, matériels et conceptuels et nous illustrerons leur usage dans certaines sciences et dans les HN.

Le modèle formel

Dans une pratique scientifique, un modèle est un artefact sémiotique servant à décrire, expliquer et comprendre des objets. Il est dit formel si les formes sémiotiques qu'il contient constituent un ensemble de symboles strictement régis par des règles que, par commodité, nous appelons souvent *langues formelles*¹² et où les symboles de départ constituent un *vocabulaire* et les règles définissent une *syntaxe*. Ce langage est sans sémantique. Cependant, un modèle formel est dit interprété si, dans une application à un domaine spécifique, une sémantique lui est ajoutée. On distingue des modèles formels selon le type de systèmes symboliques qu'ils utilisent. Dans le cadre de l'analyse reliée aux HN, nous nous concentrerons sur trois types dont deux se retrouvent souvent au sein des HN, soit les modèles *axiomatiques*, mathématiques et/ou *computationnels*.

Les modèles formels axiomatiques

Les modèles formels de type axiomatiques utilisent des systèmes symboliques dont la syntaxe contient des règles spécifiques pour la génération et la transformation des formules (inférence, substitution, etc.). De plus, certaines formules sont choisies comme postulats, axiomes, lemmes, corollaires, etc. Ces caractéristiques rendent les systèmes axiomatiques très utiles pour la construction de séquences de formules servant de démonstration et de preuve. Les prototypes de systèmes axiomatiques

¹² En anglais, on utilise le terme *language* pour désigner la langue bien que, en français, le terme *langage* désigne plus que celle-ci. Ce dernier terme est lui-même problématique car il existe des systèmes symboliques qui ne sont pas des langues au sens strict, par exemple, les systèmes symboliques mathématiques ou encore des systèmes symboliques contenant des formes graphiques comme $\Rightarrow \dot{E} + *$, des lignes, des cercles, des figures (graphes, plans, etc.).

sont les divers systèmes logiques : logiques propositionnelle, prédicative, combinatoire, etc. Les systèmes axiomatiques ne reçoivent que rarement une sémantique. On les utilise surtout pour modéliser des structures mathématiques. De fait, peu de sciences les utilisent pour la modélisation de leur objet de recherche.

Les modèles formels mathématiques

Les modèles formels de type mathématique utilisent aussi des systèmes symboliques formels : ceux-ci se caractérisent notamment par l'inclusion des classes de symboles (constantes, variables, opérations, etc.). On leur ajoute des types particuliers de règles pertinents pour des manipulations complexes de leurs symboles (substitution, élimination, récursion, transformation, etc.). Dans la majorité de leur utilisation, ils ne définissent qu'implicitement leurs axiomes ou postulats¹³. Cependant, certaines formules sont privilégiées, plus particulièrement celles exprimant des relations fonctionnelles. Les prototypes de ces systèmes symboliques mathématiques sont l'arithmétique, les algèbres, les statistiques, la géométrie, le calcul lambda, les théories des ensembles, des groupes, des graphes, du chaos, les grammaires catégorielles, etc. Par définition, ces systèmes symboliques sont dépourvus de sémantique mais, dans leur utilisation concrète à des fins scientifiques, ils reçoivent une sémantique spécifique construisant ainsi un modèle formel mathématique. Ce modèle relie le système symbolique à divers types d'entités et de relations du domaine étudié. Ces modèles formels mathématiques sont omniprésents tant dans les sciences naturelles que dans les SHS.

Les modèles formels computationnels

Les modèles formels de type computationnel sont un sous-ensemble des modèles formels mathématiques. Comme ces derniers, ils utilisent des systèmes symboliques qui se caractérisent par un type spécifique de manipulations de symboles et de formules. En effet, ces systèmes permettent d'effectuer un calcul sur les formules et de décider si elles appartiennent ou non au système dans lequel elles sont définies. Si tel est le cas, ces formules sont alors dites *calculables* et *décidables*¹⁴.

Comme le soulignait Kurt Gödel, l'une des grandes contributions d'Alan M. Turing (1937) fut que cette propriété de décidabilité et calculabilité est équivalente à des procédures effectives de manipulation de symboles réalisées par un type particulier de *machine physique* (appelée plus tard « machine de Turing »). On dit alors que cette machine « comptait » cette formule. En d'autres termes, Alan M.

¹³ On présente ces systèmes axiomatiques de manière explicite surtout en mathématique fondamentale. Par exemple, G. Peano le fait pour l'arithmétique et H. B. Curry et R. Feys (1958) pour la logique combinatoire.

¹⁴ Si toutes les formules du système sont considérées comme calculables et décidables, alors celui-ci est dit complet.

Turing montrait que *calculer* n'était pas uniquement une opération « formelle », mais aussi une opération mécanique réalisable par une machine physique, c'est-à-dire par la « computation » que faisait un « *computer* » ou, dans les termes d'Alan M. Turing, un « *calculator* » mécanique¹⁵.

Plus tard, la thèse de Church-Turing affirme que des formules calculables et computables par une machine de Turing sont équivalentes¹⁶ à d'autres formulations de systèmes symboliques mathématiques (moins mécaniques) tels le lambda-calcul d'Alonzo Church (1937), la logique combinatoire (Curry, Feys, 1958) les algorithmes (Markov, 1960), les règles de production (Post, 1936) les grammaires de Noam Chomsky et Marcel Paul Schützenberger (1963) et bien d'autres. En résumé, si une formule d'un système symbolique est décidable, elle peut être traduite par un algorithme dont une forme d'écriture est un langage de programmation. Il y a donc un lien étroit entre les systèmes symboliques de type mathématique et ceux de type computationnel. Si un système mathématique assure la calculabilité et la décidabilité de ses formules alors il est aussi un système formel computable.

Tout comme dans les systèmes axiomatiques et les systèmes mathématiques, les systèmes formels computationnels peuvent recevoir une sémantique. Cette sémantique relie les symboles et les formules aux divers types d'entités et de relations retenus du domaine étudié. La computationnalité est importante, car toutes les fonctions mathématiques n'ont pas cette propriété. Celles computables forment un ensemble éminemment restreint des fonctions mathématiques, la grande majorité des fonctions mathématiques étant non computables¹⁷. Cette notion de non-computabilité vient donc déstabiliser la construction d'un modèle formel mathématique. Car tous ne permettent pas la création de modèles computables. Pour illustrer ces concepts de modèle mathématique, nous prenons un exemple simple, mais que nous utilisons plusieurs fois dans l'analyse, soit un modèle formel qui contient la formule algébrique suivante générée par une grammaire de type algébrique : $F = Ma$.

Les règles de l'algèbre élémentaire permettent de manipuler ces symboles et de les transformer en formules suivantes : $a = M/F$ ou $M = F/a$. Selon les règles de cette grammaire, il est possible de substituer des chiffres à ces formules. Par exemple $F = 15$, $M = 5$, et $a = 3$. Comme on le sait, par ailleurs, cette formule $F = Ma$ est

¹⁵ On appelait en anglais *computors* ou *calculators* les personnes qui effectuaient concrètement le calcul. Un calculeur (*calculator*) humain, pourvu d'un stylo, de papier et d'instructions explicites, peut être considéré comme un avatar de machine de Turing (« *a human calculator, provided with pencil and paper and explicit instructions, can be regarded as a kind of Turing machine* », Church, 1937 : 42-43). Pour autant qu'on le sache, c'est ce que J. von Neuman appelle *computing machine* ou *computer*.

¹⁶ La thèse Church-Turing établit une équivalence formelle entre des fonctions computationnelles, les fonctions récursives, les algorithmes et les machines de Turing : « Les mêmes classes » de fonctions partielles et totales peuvent être obtenues dans chaque cas » (« *The same class of partial functions and of total functions can be obtained in each case* », Rogers, 1967 : 39).

¹⁷ Comme l'ont montré plusieurs mathématiciens et informaticiens, et particulièrement G. J. Chaitin (1998), la non-décidabilité est partout dans ces systèmes formels.

calculable. Ainsi, si on lui substitue ces valeurs, alors les trois formules précédentes sont transformées en : $F = Ma$ soit $15 = 5 \times 3$; $a = M/F$ soit $3 = 15/5$; $M = F/a$ soit $5 = 15/3$. Certes, l'exemple est simple, mais il met en évidence a) que le système algébrique ne manipule que des symboles $a, M, F, =, /$ qui n'ont pas encore de signification et b) que, pour effectuer des calculs, il faut effectuer au moins deux actions de manipulation : 1) une substitution des lettres par des symboles numériques et 2) appliquer des opérations mathématiques sur des chiffres. On pourrait aussi décrire ces dernières actions dans un langage formel computationnel *via* des algorithmes et un programme.

Ainsi la formulation algébrique $F = Ma$ et sa traduction algorithmique sont-elles des expressions engendrées par un langage formel et sans sémantique. Cependant, leur utilisation en physique mécanique newtonienne leur attribue une sémantique ; ces termes renvoient à la masse (M), l'accélération (a) et la force (F). Ce n'est qu'à ce moment de leur utilisation dans un modèle qu'ils contribuent à la compréhension d'un phénomène particulier.

Les modèles formels dans les sciences

La philosophie des sciences d'influence positiviste (par exemple, Rudolf Carnap, Carl Gustav Hempel, Patrick Suppes) a proposé que les systèmes formels axiomatiques constituent la forme idéale de l'explication scientifique. Mais cette proposition n'a pas vraiment résisté aux critiques épistémologiques. Des études empiriques de pratiques scientifiques ainsi que de nombreux arguments théoriques (Winther, 2015) ont montré que peu de pratiques scientifiques utilisent effectivement des modèles formels de type axiomatique.

Pour leur part, les modèles mathématiques ont mis en évidence les problèmes que pose la dissociation intrinsèque de la syntaxe et de la sémantique pour un langage formel. Lorsqu'ils sont appliqués à des phénomènes d'une grande complexité, d'une forte irrégularité, d'une puissante dynamique, tels le climat, les formes économiques, l'environnement, les mouvements sociaux, les grandes entreprises, le cerveau, la culture, les grands projets d'ingénierie civile, etc., ils utilisent des formes ou des structures mathématiques d'une grande complexité où domine souvent la non-linéarité¹⁸ telles les systèmes mathématiques probabilistes, stochastiques, dynamiques, chaotiques, évolutionnistes, morphodynamiques, etc.

Par exemple, on retrouve ces types de modèles formels en économie qui, pour expliquer des dynamiques de marché, en appelle à la théorie des jeux, aux systèmes d'équilibre, au calcul d'optimisation, aux probabilités bayésiennes, etc. Les

¹⁸ La non-linéarité désigne la propriété de fonctions mathématiques dont les sorties ou valeurs ne sont pas proportionnelles aux données entrantes. Dans ce cas, la représentation de la variation de la fonction ne peut être représentée par une ligne droite.

neurosciences explorent des modèles formels d'algèbre non linéaire de classification, de catégorisation et qu'elles nomment métaphoriquement des réseaux de neurones. L'archéologie contemporaine utilise simultanément ou séquentiellement des modèles statistiques, optiques, chimiques, physiques, linguistiques qui, ensemble, participent à l'explication d'une fouille. Et les chercheurs qui utilisent ces modèles mathématiques et computationnels connaissent les contraintes et limites qu'ils imposent à une description et une explication scientifiques.

En effet, tant pour la construction que pour le calcul effectif, ces systèmes formels présentent de nombreuses contraintes qui ne sont pas sans affecter la qualité de leur utilisation en sciences. On doit souvent atomiser ou discrétiser des phénomènes (même si certains langages mathématiques permettent le traitement du continu) risquant ainsi de ne pas représenter correctement les conditions initiales ; on doit faire un choix dans les propriétés et les facteurs à retenir réduisant ainsi la qualité de la description et de l'explication ; et, pour computer effectivement ces types de fonction, on peut réclamer un traitement informatique parallèle qui, lui-même, exige des paramètres de contrôles, ce qui affecte la possibilité de suivre le déroulement du traitement, c'est-à-dire sa traçabilité. Voilà autant de contraintes qui, dans de nombreuses recherches, entraînent des problèmes de précision, de prédiction, de certitude, de validité, de complétude et donc de non-calculabilité, etc. :

« La non-décidabilité est l'incomplétude sont partout, des mathématiques aux sciences de l'informatique, à la physique, aux sections mathématiquement formalisées de la chimie, de la biologie, de l'écologie et de l'économie » (Chaitin, Doria, Costa, 2012 : 2)¹⁹.

Enfin, dans certains cas, cette complexité est telle qu'il devient difficile de privilégier une fonction mathématique particulière, parce qu'elle n'est pas encore connue. Elle se transforme en l'objet de la recherche, c'est-à-dire que le but de la recherche sera d'identifier cette fonction²⁰. Aussi préfère-t-on utiliser des stratégies inductives pour explorer des régularités, découvrir des invariants et les exprimer par des équations. Les modèles computationnels d'apprentissage machine et d'apprentissage profond²¹ appliqués à des données massives permettent des stratégies d'analyses inductives.

Ainsi, même si ces modèles mathématiques et computationnels sont puissants, ne conviennent-ils pas nécessairement à toutes les démarches scientifiques. Qu'en est-il des phénomènes sémiotiques qui intéressent les HN ?

¹⁹ « *Undecidability and incompleteness are everywhere, from mathematics to computer science, to physics, to mathematically-formulated portions of chemistry, biology, ecology, and economics* ».

²⁰ Par exemple, on observe les variations de la bourse. Mais, chanceux celui qui découvrirait la fonction mathématique qui permettrait de déterminer les valeurs à venir.

²¹ L'apprentissage machine et l'apprentissage profond sont des techniques algorithmiques sophistiquées qui, à partir d'échantillons aléatoires de données descriptives, visent à approximer la fonction générale expliquant ces données (Mitchell, 1997 ; Schmidhuber, 2015 ; Goodfellow, Bengio, Courville, 2015). On dit ainsi que ces algorithmes « apprennent » à généraliser à partir de cas particuliers.

Les modèles formels dans les humanités numériques

Selon une première vision, le domaine classique des humanités ne semble pas devoir se préoccuper de modélisation formelle, si ce n'est peut-être pour certains linguistes ou anthropologues structuralistes qui en ont fait une norme idéale de la démarche. Mais cela change lorsque le territoire des humanités fusionne avec celui des *humanities* et intègre l'ordinateur : les modèles formels prennent soudainement une grande importance.

Comment peut-on comprendre le rôle des modèles mathématiques et, surtout, des modèles computationnels au sein de cette discipline naissante ? Avant de répondre à cette question, il faut préciser que ce terme *numérique* n'est pas adéquat pour comprendre la problématique de la computation dans les humanités « numériques ». Au sens strict, cette propriété numérique n'explique aucunement le rôle de l'informatique dans ce territoire. Ce n'est pas parce qu'une pratique utilise des symboles numériques de types binaires que leur modélisation peut être dite computationnelle. Comme nous l'avons soutenu ailleurs (Meunier, 2014), une telle définition est problématique et masque leur nature véritable. En effet, on ne met pas en évidence que, derrière ces outils, se construit une modélisation formelle qui repose sur l'un ou l'autre des systèmes symboliques mathématiques et computationnels.

L'utilisation des ordinateurs dans les HN est autrement plus complexe. En effet, l'ordinateur a permis aux chercheurs de s'approprier des modèles computationnels et, ultimement, les modèles mathématiques d'une grande puissance et fécondité pour les assister dans leurs tâches classiques d'interprétation.

Certains des logiciels utilisés reposent sur des modèles mathématiques et computationnels relativement simples (grammaires, règles, arithmétiques, statistiques élémentaires, etc.). Voici un exemple simple : en lexicométrie, on trouve des programmes qui permettent de calculer un rapport entre la fréquence des lexèmes et leur rang. Ce programme repose sur un modèle formel qui contient l'énoncé/équation suivante :

$$f(n) = \frac{K}{n}$$

Sur le plan mathématique, cette équation ne fait que représenter une relation fonctionnelle entre des nombres. Mais, lorsque les symboles désignent des valeurs numériques associées à des ensembles de mots dans un texte, l'équation reçoit le nom « loi de Zipf ». Elle décrit la *fréquence* des mots $f(n)$ en regard du *rang* n de leurs fréquences (K est une constante). Ainsi, si le mot le plus fréquent (rang 1) est de 10 000 alors la fréquence de rang 10 sera de 1 000 et ainsi de suite.

Par ailleurs, on sait que cette équation mathématique est calculable et donc qu'elle peut recevoir une traduction en algorithmique ou dans un programme²². Cependant, si cette équation n'est pas computable, la traduction algorithmique est impossible à moins de l'assister d'heuristiques. Autrement dit, il n'y a pas d'humanités dites numériques s'il n'y a pas de modélisation formelle contenant des propositions symboliques mathématiques ou équations ou règles calculables traduisibles en algorithmes. Ainsi, parce qu'utilisant l'ordinateur, tout projet sérieux en HN met-il – implicitement ou explicitement – en place une modélisation formelle.

Il existe une multiplicité d'exemples d'ensembles d'algorithmes réunis dans des logiciels et utilisés dans les HN. Le concordancier en est un typique : à partir de quelques règles et calculs simples, il permet de trouver les segments des contextes d'un terme pôle. À l'aide de calculs statistiques simples, un logiciel d'indexation associe des termes significatifs (mots clés, vocabulaire contrôlé) à des documents à des fins de fouille et de rappel. Un logiciel de numérisation encode des textes pour pouvoir les traiter sur un nouveau support. À l'aide d'une grammaire, un processeur syntaxique décrit des structures de phrases. Ces logiciels sont des plus utiles pour assister les chercheurs des humanités dans des opérations qui, autrefois, étaient faites « à la main ».

Mais, de plus en plus, on voit apparaître des modèles mathématiques et computationnels complexes comme ceux qu'on trouve dans les sciences dont nous parlions plus haut telles la climatologie, les neurosciences, etc. Des classifieurs non linéaires assistent l'analyse thématique (*topic modeling*), des probabilités bayésiennes assistent la traduction, l'analyse d'images ou de la musique, etc. Des statistiques inductives de type réseaux de neurones, apprentissage machine ou apprentissage profond détectent des régularités et des invariants dans les productions sémiotiques qui sont très pertinents pour des parcours interprétatifs. Ces modèles rencontrent les mêmes limites descriptives et explicatives que lorsqu'ils sont appliqués dans d'autres domaines.

Dans les HN, ces modèles formels simples ou complexes sont devenus si transparents et bien cachés dans les logiciels qu'on en vient à oublier leur présence et tous les problèmes qui leur sont intrinsèques. On les perçoit comme des « outils » commodes et assistant les parcours interprétatifs sans se préoccuper de leur véritable nature, donc de leurs limites. C'est peut-être pour cette raison que certains chercheurs des humanités ne sont pas convaincus de la pertinence de l'utilisation de l'ordinateur dans leur domaine. Pour eux, tous les objets sémiotiques qu'interprètent les humanités ne peuvent pas être facilement soumis à une modélisation mathématique et computationnelle précisément en raison de leur complexité intrinsèque, donc de leur possible non-calculabilité. Ce qui est un argument solide. Mais ce jugement doit être nuancé, car il repose sur une conception trop restreinte de la modélisation mathématique et computationnelle. On pourrait inverser la position et considérer que la modélisation formelle est en soi une forme particulière d'interprétation, que,

²² Pour un tel programme en langage java, voir : <http://diveintodata.org/2009/09/13/zipf-distribution-generator-in-java/>.

avec d'autres formes de modélisation, elle peut contribuer à la compréhension de réalités sémiotiques, même les plus complexes. Pour ce faire, il faut comprendre la nature de la complexité de ce filtre interprétatif que sont ces modèles formels.

Les modèles matériels

Bien que contribuant à l'explication d'une réalité, un modèle formel demeure limité pour la compréhension humaine. Une théorie scientifique ne pourrait s'arrêter à la seule présentation d'une équation formelle pour décrire et expliquer un phénomène. Par exemple, un physicien ne peut expliquer le phénomène de la force en ne présentant que l'équation formelle $F = Ma$. Ainsi, outre les modèles formels, les théories scientifiques introduisent-elles d'autres modèles qui contribuent à une plus grande compréhension du phénomène étudié.

L'un de ces modèles que nous qualifions de matériel a pour fonction de repérer, parmi toutes les propriétés possibles, le fait qu'une réalité peut posséder celles qui sont pertinentes pour une explication causale et que certains appellent explication mécanique : « Les mécanicistes insistent sur le fait qu'une explication consiste à élucider les structures causales qui produisent, sous-tendent ou maintiennent le phénomène étudié » (Craver, Tabery, 2015)²³.

Ces modèles utilisent diverses formes sémiotiques pour exprimer ce rôle. Certains reprennent des formules d'une langue formelle, mais la modifient par l'utilisation de termes ou de formes iconiques techniques extraits de la langue naturelle, mais reconnus par la communauté scientifique. Ces termes et formes iconiques sont interprétés comme exprimant des relations de causalité ou des mécanismes.

Nous illustrons ce rôle d'un modèle matériel en reprenant notre exemple de l'équation $F = Ma$. Malgré son apparence algébrique, cette équation n'utilise pas des symboles issus du vocabulaire classique de l'algèbre comme le ferait l'équation $z = xy$ ou $F[x, y] : xy$. Comme on l'a vu, F , M et a sont des abréviations de termes appartenant au langage de la physique. Cette substitution n'est pas innocente ; elle repose sur une convention commune aux physiciens qui connaissent bien la sémantique de ces mots du langage naturel qui sont reliés directement à des propriétés physiques pertinentes pour une explication causale et mécanique. C'est ce qui permet à $F = Ma$ d'être interprétée comme décrivant une relation causale à partir de la masse et de l'accélération. Ainsi, s'il n'est pas naturel de dire à propos de la formule $z = xy$ que z est causé par la multiplication de x par y , et même que F est causé par M multiplié par a , il est tout à fait naturel pour un physicien de dire que *la force est causée par l'accélération d'une masse. Ou encore on explique la force par un rapport causal entre l'accélération et la masse.*

²³ « *Mechanists insist explanation is a matter of elucidating the causal structures that produce, underlie, or maintain the phenomenon of interest* ».

Dans certaines disciplines, la modélisation matérielle illustre des formules du modèle formel *via* des formes sémiotiques iconiques. Par exemple, en électricité, la loi d'Ohm est illustrée par un graphe orienté qui décrit la « mécanique » causale de la différence de potentiel de tension (U) d'un courant électrique (I) passant dans une borne de résistance (R).

Image 1. Modélisation matérielle de la loi d'Ohm



Ce type de modèle matériel peut se voir illustré par des artefacts physiques qui lui servent de prototypes ou d'instanciations. Ces artefacts permettent la compréhension des relations causales ou des mécanismes en jeu. Par exemple, des artefacts physiques de l'équation $F = Ma$ peuvent être deux boules se frappant ou une boule de démolition suspendue à une grue. On peut voir le transistor comme une instanciation de la loi d'Ohm et de son modèle graphique précédent. Ainsi, dans plusieurs démarches scientifiques, par exemple dans les sciences de l'ingénierie, de la gestion où l'expérimentation ou la construction d'artefacts est importante, ces types de modèles matériels « physiques » sont des plus importants.

De plus en plus dans les sciences contemporaines, la modélisation matérielle prend une forme particulière : elle traduit le modèle formel computationnel exprimé par des algorithmes en un programme informatique. Ce dernier reprend certes les algorithmes, mais ajoute quelque chose : des précisions sur les conditions par lesquelles un ordinateur qui, rappelons-le, n'est qu'une « machine » de type électronique²⁴, peut effectuer la computation de la fonction calculable sous-jacente à ces algorithmes. Autrement dit, ce modèle effectue la computation dans une « machina » c'est-à-dire dans des *mécanismes* physiques de type électronique. Ce qui, pour les épistémologues mécanicistes, est une manière originale de décrire et expliquer les relations causales ou les mécanismes d'un phénomène. Un modèle matériel se transforme donc et devient un modèle mécanique, mais de type informatique.

Avec le développement de cette technologie informatique, il est devenu possible d'effectuer de manière extrêmement rapide la computation de modèles formels des plus complexes et de représenter sous forme de simulations virtuelles la structure ou même l'évolution de phénomènes étudiés. Par exemple, par des techniques informatiques sophistiquées, on peut présenter sous des formes iconiques – écran deux ou trois dimensions (colonnes, grappes, cartes, dessins animés, scènes statiques ou dynamiques) – des configurations des résultats de la computation ou encore les moments de la computation. La modélisation informatique devient alors une simulation informatique.

²⁴ La machine de Turing est typiquement une machine mécanique (roue, papier) alors que l'ordinateur de type von Neumann est de type électronique.

Ces simulations sont devenues très populaires dans plusieurs disciplines, car elles ont un avantage cognitif indéniable. Parce qu'elles sont surtout analogiques, elles illustrent des relations causales et des structures mécaniques qui deviennent ainsi facilement interprétables par un humain. L'imagerie du cerveau par résonance magnétique (IRM) en est un exemple : elle ne présente pas une photographie du cerveau, mais bien une simulation informatique iconique d'aires cérébrales que des spécialistes interprètent causalement d'une façon plus facile que les modèles formels qui ont servi pour décrire la dynamique électrique d'un cerveau. Dans certains cas, des simulations sont aussi traduites dans un langage naturel métaphorique, narratif, etc. Ainsi présente-t-on des cooccurrences importantes de mots dans un texte comme des « nuages » ou des « associations » de mots, etc.

Dans cette perspective, on peut voir un robot comme un modèle physique, c'est-à-dire une simulation informatique réalisée par des mini-ordinateurs. De fait, un robot n'est qu'un assemblage hybride de mini-ordinateurs qui effectuent un calcul mécanique, c'est-à-dire qui computent des fonctions calculables modélisant des comportements moteurs et cognitifs. Il n'est pas vrai qu'un robot « pense », « agi », « décide », « voit », etc. Ce sont des termes anthropomorphiques utilisés pour nommer des formes de visualisations 3D, instanciation de fonctions computables. Dans cette perspective, un robot est une forme spectaculaire de modélisation informatique de processus cognitifs.

La modélisation matérielle dans les sciences

Des chercheurs (Bäetu, 2013 ; Rheinberger, 1997 ; Leonelli, 2007 ; Weisberg, 2007, 2015 ; Armatte, Dahan Dalmedico, 2004) ont montré que la modélisation matérielle, surtout informatique, est devenue omniprésente dans les sciences contemporaines où elle joue un rôle cognitif important. Plusieurs sciences naturelles – par exemple la climatologie, l'économie, l'ingénierie l'astrophysique, les sciences de l'environnement, la médecine, etc. – sont confrontées à des phénomènes difficiles à observer dans un temps court et dont certaines propriétés sont bruitées, irrégulières, évolutives. Comme nous l'avons souligné (voir *supra*), ces sciences ont recours à des modèles formels d'une grande complexité mathématique. On peut penser aux modèles statistiques inductifs, aux modèles dynamiques, chaotiques, stochastiques probabilistes, neuronaux, etc. Or, malgré leur complexité, ces modèles mathématiques ont pu être utilisés parce qu'ils pouvaient être calculés concrètement par un ordinateur. Ce qui était auparavant impossible.

En sciences cognitives, où l'objet d'étude n'est pas une réalité que l'on peut facilement observer et expérimenter, la modélisation informatique est devenue un outil des plus heuristiques. Elle a permis aux modèles formels connexionnistes²⁵ d'être explorés sous diverses facettes. Les neurosciences ne seraient pas ce qu'elles sont aujourd'hui sans l'IRM.

²⁵ La modélisation connexionniste est utilisée dans diverses sciences cognitives. Elle modélise les opérations cognitives comme des ensembles de connexions entre des neurones, c'est-à-dire comme

Malgré ce succès de la modélisation informatique, toutes les sciences ne l'intègrent pas dans leur démarche. Par exemple, certaines recherches en sociologie ont des difficultés à modéliser les représentations sociales *via*, par exemple, des simulations informatiques. Ce ne sont pas tous les domaines de la psychologie qui acceptent facilement de décrire et expliquer les émotions, la conscience *via* ces types de modèles. La science politique n'explique que très peu les relations politiques entre pays par des simulations informatiques. Enfin, tous les chercheurs ne sont pas d'accord pour interpréter ces types de modèles physiques matériels comme des descriptions de relations causales ou de mécanismes.

Malgré ces difficultés, il est important de comprendre ce type de modèle matériel car, lorsque les sciences introduisent l'ordinateur dans leur démarche, elles sont nécessairement confrontées à la question de la modélisation informatique, c'est-à-dire de la présentation des conditions matérielles d'une computation effective. Même si une science possède des modèles formels des plus complexes et des traductions algorithmiques élégantes et complètes, elle ne possède pas nécessairement du même coup un modèle de leur implémentation dans un ordinateur concret qui possède une structure matérielle puissante pour réaliser le calcul effectif : disque dur, mémoire, horloge, etc.

Les modèles matériels informatiques dans les humanités numériques

Le modèle informatique est évidemment des plus importants dans les sciences contemporaines, mais, dans les humanités dites numériques, il l'est davantage. Il en délimite le territoire :

« Les *humanités computationnelles* sont précisément l'automatisation de toutes les analyses des expressions de l'humain (donc une activité parfaitement « humaniste ») dans le sens le plus large de cette expression et vont de la musique au théâtre, du design et de la peinture à la phonétique, mais dont le noyau demeure le discours dans les textes écrits » (nous soulignons, Busa, 2008 : 8)²⁶.

En effet, cette modélisation informatique joue un rôle épistémique important dans les humanités numériques. Ainsi permet-elle de traiter des données sémiotiques d'une ampleur nouvelle, d'utiliser des modèles formels et computationnels originaux. Placée aux bons endroits dans la démarche de recherche, elle participe au processus interprétatif de manière originale. De ce fait, elle en modifie les objets et la méthode et introduit donc des angles nouveaux pour leur compréhension.

des réseaux de neurones. Ces derniers sont eux-mêmes formalisés *via* une algèbre matricielle. On les décrira comme des neurones artificiels.

²⁶ « *Humanities computing is precisely the automation of every possible analysis of human expression (therefore, it is exquisitely a 'humanistic' activity), in the widest sense of the word, from music to the theater, from design and painting to phonetics, but whose nucleus remains the discourse of written texts* » .

Dans les multiples projets des HN, on trouve une multiplicité de programmes et de logiciels. Quelques fois, certains ont été construits par des utilisateurs aguerris aux techniques de programmation ; d'autres seront livrés clé en main et disponibles sur des sites ou sur le marché spécialisé (par exemple, la boîte à outils TAPOR). Quelques fois encore, ce sont des outils autonomes ou propriétaires et même commerciaux (*Alceste*, *Sato*, *Sphynx*, *Antconc*, *Rapid Miner*, *Provalis*, etc.). D'autres sont des modules préprogrammés (applications) intégrables dans des langages de programmation comme R, Python, etc. Enfin, d'autres sont construits sur mesure par le chercheur. Mais il faut toujours rappeler que ces logiciels permettent de computer des algorithmes qui expriment, dans un langage de programmation, les fonctions computables des modèles formels.

Certains des modèles formels d'arrière-plan, tels la loi de Zipf, sont relativement simples à comprendre, parce qu'exprimés dans un langage algébrique bien connu des chercheurs, mais d'autres, comme nous l'avons vu plus haut, tels les modèles linéaires et non linéaires de classification, d'optimisation, de distributions statistiques et même de grammaires formelles, etc., sont souvent très complexes, et donc difficiles à comprendre par les chercheurs des humanités.

Un des grands dangers des humanités numériques est que les chercheurs soient séduits par une technologie informatique facilement accessible, mais dont ils ne comprennent que minimalement les modèles formels mathématiques et computationnels. Pensons ici aux logiciels d'analyse de textes comme *Text Miner*, *Alceste*, *Antconc*, etc. ou aux applications de « R » pour l'analyse thématique (*topic modelling*). Ceux-ci reposent sur les modèles formels complexes avec des choix de paramètres spécifiques. Dans certains cas, leur utilisation aveugle peut être pertinente et heuristique, mais dans bien d'autres cas, elle risque de produire des résultats dont l'interprétation est fautive parce que dissociée des algorithmes et des paramètres qui ont généré les résultats. À l'inverse, on risque aussi de dénaturer le phénomène sémiotique étudié en ne le décrivant, l'expliquant et le comprenant que sous l'angle des algorithmes et paramètres du logiciel utilisé.

Un exemple de ceci est le nuage de mots. Des algorithmes traitent des mots et ils produisent des représentations graphiques qui iconisent des relations entre des mots. Ils sont des simulations des associations entre des occurrences de lexèmes dans un corpus. Mais l'interprétation qui peut être appliquée à ces nuages peut être radicalement différente de ce que les algorithmes permettent. Aussi peut-il arriver que l'on utilise des outils informatiques et que les résultats effectivement produits reposent sur un modèle formel qui, ultimement, ne soit pas des plus pertinents pour décrire, expliquer et comprendre l'objet sémiotique que s'est donné un projet des humanités numériques.

Ainsi pourrait-on se demander si les résumés automatiques, très utiles dans certains cas, sont pertinents lorsqu'appliqués à des œuvres littéraires. De tels programmes pourraient résumer à 10 %, 40 % 75 % d'*À la recherche temps perdu* de Marcel Proust. Mais que vaudrait ce résumé pour un critique littéraire ?

Les chercheurs des humanités numériques ne doivent utiliser des outils informatiques pour leur recherche que s'ils comprennent minimalement le modèle formel qui sous-tend ces outils. Sinon, leur recherche court le danger de parvenir à des résultats superficiels qui n'assistent que très peu la description, l'explication et la compréhension des artefacts sémiotiques étudiés. Mais une compréhension sérieuse des modèles formels computationnels et informatiques peut, lorsqu'ils sont placés adéquatement dans la démarche interprétative, devenir des outils importants et originaux d'exploration de domaines sémiotiques.

Le modèle conceptuel

Une théorie scientifique ne peut se contenter de présenter des modèles formels et matériels. Car ceux-ci demeurent limités pour la compréhension humaine. Même si l'on présentait à un sujet deux boules qui se frappent, que l'on expliquait la force qui les traverse uniquement par une équation formelle comme $F = Ma$, et que l'on l'illustrait ceci par une simulation informatique, la compréhension n'en demeurerait pas moins pauvre et limitée. On ajoute donc, habituellement, un troisième type de modèle qui permet aux humains une intelligibilité de la réalité, mais qui est plus intégrée à leur univers conceptuel général.

Tout comme les autres modèles, celui-ci, que nous appelons *conceptuel*, est aussi constitué d'énoncés. Mais, cette fois, ceux-ci sont exprimés dans une langue naturelle. Du point de vue d'une sémantique extensionnelle, les termes conceptuels utilisés réfèrent à des entités et des propriétés de la réalité. Au début de la recherche, ces énoncés visent à exprimer des intuitions ou des croyances à propos des propriétés de la réalité étudiée. Ils sont considérés comme des hypothèses qui ont à être validées en regard des modèles formels et matériels. Souvent, de multiples itérations et révisions sont nécessaires. Finalement, lorsque la démarche a atteint une certaine maturité et que sa démarche et ses résultats ont été acceptés par une communauté épistémique, ces énoncés sont considérés comme des lois ou des croyances justifiées... jusqu'à un prochain ajustement, contre-exemple, etc.

Du point de vue d'une sémantique intensionnelle, la signification des énoncés du modèle conceptuel est cependant plus complexe. Elle est souvent de type holiste. La conceptualisation qui y est exprimée ne dépend pas uniquement du lexique, de l'énoncé ou de l'ensemble de la théorie prise isolément. Elle dépend aussi de diverses structures externes de compréhension implicitement ou explicitement associées à la théorie construite. Dans les théories cognitives contemporaines, certains appellent ces structures « espace conceptuel » (Gärdenfors, 2000), « système conceptuel » (Brown, 2007), « modèles mentaux » (Johnson-Laird, 1983), « connaissances conceptuelles tacites » (Polanyi, 1967) ou même, lorsque partagées socialement, des « paradigmes » (Kuhn, 1962) ou des « représentations sociales » (Jodelet, 2003). On peut interpréter ces univers comme des variantes, des

précisions ou des dimensions conceptuelles de ce que la tradition herméneutique peut inclure dans des visions du monde ou des *Weltanschauung*. Enfin, sur le plan pragmatique, ce type de modèle conceptuel sert à la communication, la discussion, la révision, l'ajustement et la négociation des concepts et il contribue à construire des communautés épistémiques.

Dans une telle perspective, le modèle conceptuel, comme les deux autres modèles retenus, joue un rôle très important de médiation entre l'objet de recherche et la compréhension humaine. Il est le lieu de l'intelligibilité ultime. Sans lui, les modèles formels et informatiques sont assurément utiles pour la manipulation des symboles par des machines, mais leur compréhension exige ultimement une forte relation avec un modèle conceptuel.

Mais parce qu'intimement lié au langage naturel, le modèle conceptuel souffre de toutes les lacunes et imperfections propres à ce langage tant sur le plan syntaxique et sémantique que sur le plan pragmatique (ambiguïté, instabilité aléthique, actes de langage, etc.). Il profite aussi de sa richesse et de sa créativité (narration fiction, métaphore, rhétorique, etc.). Malheureusement, comme ce type de modèle est naturel et transparent, on en oublie sa présence, et les théories de la science n'en parlent pas.

Les modèles conceptuels dans les sciences

Ce type de modèle conceptuel est omniprésent en sciences. Une illustration classique est la théorie copernicienne. Dans ses origines, cette théorie était constituée de modèles formels mathématiques décrivant et prédisant le mouvement des planètes. Son modèle matériel prenait la forme de graphes circulaires. Le modèle conceptuel explicitait certains énoncés du modèle formel, mais les reliait à l'espace/cadre/système conceptuel de l'astronomie et de la religion de l'époque. Les historiens et épistémologues ont montré que l'objet des débats importants sinon même virulents que la théorie copernicienne a provoqué relevait avant tout du modèle conceptuel par lequel on interprétait ces modèles formels et ces illustrations. En effet, il mettait à l'épreuve l'espace conceptuel concurrent lié à l'expérience phénoménologique du mouvement de la Terre par rapport au soleil. De plus, il était confronté à l'espace conceptuel religieux sur l'origine divine de ce mouvement. Il brisait l'intelligibilité ultime que la société de l'époque avait de ce phénomène.

Une autre illustration est plus contemporaine. Elle vient des neurosciences. Nous la tirons du livre de grande qualité de Michael S. A. Graziano, *Consciousness and the Social Brain* (2013). Ce dernier contient un ensemble de modèles formels et matériels du fonctionnement conscient du cerveau. Voici un extrait de son introduction :

« Le cerveau est composé de neurones qui se transmettent de l'information entre eux. L'information est liée plus efficacement entre chaque neurone et maintenue encore plus efficacement sur une

longue période si les signaux électriques des neurones oscillent de manière synchronisée. En conséquence, la conscience peut être causée par l'activité électrique de plusieurs neurones oscillant ensemble » (nous soulignons, Graziano, 2013 : 6)²⁷.

Nous ne trouvons pas ici d'équation, ni de programme informatique. Nous lisons les énoncés en langue anglaise qui expriment des conceptualisations issues des sciences cognitives passées ou contemporaines : par exemple, neurosciences, psychologie de l'apprentissage, de la théorie de la communication et de l'information, de la philosophie de l'esprit, etc. Et on trouve des énoncés qui décrivent des neurones qui *communiquent, se parlent, échangent de l'information, etc.* et affirment que le cerveau *pense, mémorise, est conscient, etc.* Ces types d'énoncés occupent près de 80 % du livre ! Et ils sont essentiels à la compréhension des modèles formels, des programmes présentés dans la théorie défendue. C'est par eux que l'auteur décrit et explique la « conscience » et qu'il en suggère une intelligibilité en la reliant à nos connaissances générales.

Quelle que soit la science que l'on peut prendre comme exemple, aucune n'échappe à cette modélisation conceptuelle effectuée dans une langue naturelle. On peut espérer contraindre ce modèle en le réduisant à un métalangage formel (Tarski, 1933), mais aucune science, y compris mathématique, n'a réussi à éliminer complètement une forme ou l'autre de modélisation conceptuelle exprimée en langue naturelle. Même un programme informatique doit avoir sa « documentation », sinon il est incompréhensible pour un humain. Albert Einstein a lui-même traduit ses modèles formels en termes de théorie de la « relativité ». Kurt Gödel a présenté son théorème en termes « d'incomplétude ». Burrhus F. Skinner a traduit des lois du « *behavior* » en termes de « conditionnement opérant ». Edward Lorenz a parlé de son modèle chaotique en termes d'« attracteur ». Et Allen Newell et Herbert A. Simon ont expliqué leurs programmes informatiques comme réalisant de l'« intelligence artificielle ». James L. McClelland et David E. Rumelhart ont traduit leurs modèles vectoriels d'optimisation en « *réseau de neurones* », sinon même en « *mind* ». Les climatologues finissent toujours par dire que les états des stabilisations de multiples facteurs climatiques de leur modèle formel permettent de prévoir que « demain il y aura des averses de pluie ».

En résumé, un modèle conceptuel présente dans un langage naturel une conceptualisation du phénomène étudié, mais tel que le chercheur la conçoit, peut l'exprimer et la communiquer. Une science ne peut expliquer des réalités que si, quelque part dans sa démarche, elle construit un modèle conceptuel de son objet de recherche, si limité, ambigu, peu rigoureux soit-il. Certains de ces modèles conceptuels se divisent en de multiples sous-modèles qui se concentreront sur l'un ou l'autre sous-territoire de l'espace du problème étudié.

²⁷ « *The brain is composed of neurons that pass information among each other. Information is more efficiently linked from one neuron to another, and more efficiently maintained over short periods of time, if the electrical signals of neurons oscillate in synchrony. Therefore consciousness might be caused by electrical activity of many neurons oscillating together* ».

Certaines critiques venant surtout des positions épistémologiques positivistes et naturalistes affirment que ce type de modèle n'exprime que des observations superficielles et se résume à des descriptions superficielles dépendantes d'une culture ou d'attitudes psychologiques populaires. Il apparaît comme ambigu et intuitif, chargé d'intentions et d'affects, fondés sur des observations naïves et des analyses phénoménologiques immédiates dont la valeur « empirique » est faible et discutable. Il pourrait être mieux structuré et contrôlé dans son expression et devrait en appeler à de meilleurs outils d'observation. Pour ces raisons, ils le rejettent.

Certes, on peut rêver réduire toute l'explication à un modèle formel confirmé par des expérimentations ou des simulations, mais, de ce fait, on élimine un intermédiaire, c'est-à-dire un médiateur essentiel à la compréhension. Une théorie scientifique qui ne contient que des équations et des artefacts les implémentant sera difficile à comprendre... par un humain. Reste que, en raison de ces problèmes, il faut utiliser avec prudence la modélisation conceptuelle, car elle demeure cognitivement fragile. Ce n'est souvent qu'en relation aux autres modèles et aux nombreux autres espaces conceptuels que ses énoncés finissent par participer de manière constructive à une théorie.

Enfin, ce modèle conceptuel est souvent si transparent pour un chercheur qu'il en vient à être oublié. Ou, si on l'accepte, on ne le voit pas comme *modèle*. La construction de ce modèle ne sera pas toujours des plus faciles. Le monde ne saute pas de manière spontanée dans des modèles ! D'une part, il dépend des capacités cognitives humaines (observation, intuition, expérimentations directes ou assistées de technologies) ; d'autre part, il est sensible aux négociations d'un consensus langagier entre les membres d'une communauté épistémique et aux biais épistémiques qui le traverseront. On y trouve des croyances dont les origines relèvent de l'expérience subjective, de la culture générale ou de la fonction sociale, etc. Autrement dit, c'est ici que l'herméneutique a son emprise. Une approche pragmatique de la science qui voit une théorie comme construction de modèles met en évidence le fait que la science est un processus complexe d'interprétation. Dans ce processus, la modélisation conceptuelle met la théorie en relation avec la culture propre au chercheur et sa communauté épistémique n'y est pas étrangère. Malgré toutes ses limites, une théorie scientifique ne peut se passer d'un modèle conceptuel.

La modélisation conceptuelle dans les HN

Les HN en appellent aussi à une modélisation conceptuelle. Et comme dans les autres modélisations, elle est omniprésente. En effet, aucun projet des HN ne peut l'éliminer; et ceci en raison de la sémiotique de leurs multiples objets. La modélisation conceptuelle qui les accompagne est inévitable malgré la force des modèles formels et la puissance des ordinateurs.

Prenons pour exemple concret le projet de numérisation des écrits alchimiques d'Isaac Newton.²⁸ Il a fallu presque 15 ans pour conceptualiser les diverses dimensions de cet objet sémiotique. Il a fallu construire des modèles conceptuels très complexes pour traiter les critères d'authenticité des manuscrits, les divers types de symboles semi-iconiques utilisés dans les énoncés alchimiques, les transcriptions diplomatiques et standard, les formes de fouille adaptées à ce type d'objet sémiotique, le contenu du site web répondant à l'utilisateur débutant ou expert, etc. Toutes ces opérations impliquaient une multitude de modèles s'exprimant sous forme de textes, d'icônes, etc. La modélisation formelle et les algorithmes de numérisation, d'indexation, d'annotation, de classification, etc. sont relativement simples par rapport au modèle conceptuel requis pour réaliser ce projet.

Prenons encore ici l'exemple des visualisations dites « nuage de mots » où certains mots sont mis en évidence par leur position, leur forme ou leur couleur. Un nuage de mots n'est que l'expression iconique de traitements informatiques appliquant certaines fonctions mathématiques computationnelles sur des mots. Ce sont des sacs ou des cohortes de mots dont on a précisé certaines propriétés. Une interprétation rigoureuse ne peut que relier la couleur et/ou la grosseur des mots à ces fonctions. Mais une telle interprétation est peu intéressante. Presque inévitablement, on en arrive à interpréter ces nuages en les situant dans des champs conceptuels périphériques qui tous relèvent de théories sémantiques, psychologiques ou philosophiques du sens ou du concept. Ainsi les mots seront-ils vus tantôt comme des « thèmes dominants » (théorie littéraire), des « connaissances » (intelligence artificielle), des thesaurus conceptuels (science de l'informatique), des concepts importants (philosophie), des « noyaux » (sociologie des représentations sociales), des arguments (rhétorique), des espaces mentaux (philosophie de l'esprit), etc. Imaginons une étude des discours de Donald Trump dont le nuage de mots contiendrait en ordre de décroissance d'importance visuelle, les mots *femmes*, *Mexique*, *immigration* et *mur*. On peut deviner la conceptualisation de l'interprétation qui serait faite. Et on peut s'interroger sur le cadre théorique ou même culturel qui la fonde.

Notre recherche sur l'analyse conceptuelle assistée par ordinateur (Meunier, Forest, 2009 ; Chartrand et al., 2016) a exigé de nombreuses clarifications conceptuelles pour comprendre plus clairement les principes d'un tel type d'analyse. Ce n'est qu'après cette conceptualisation que des modèles formels et informatiques ont pu être choisis. Cette modélisation conceptuelle fut une partie importante de notre démarche. Malgré des modèles formels sophistiqués issus des sciences cognitives et des algorithmes complexes qui les traduisaient, et un modèle informatique solide, le défi le plus profond fut de conceptualiser le rapport entre les multiples expressions linguistiques et le concept de « *mind* » chez Charles S. Peirce. Le contenu conceptuel de ce mot se réduit-il au sens, à la signification du mot anglais *mind* ? Comment ce mot opère-t-il dans l'ensemble des textes de Charles S. Peirce pour révéler les propriétés de ce concept ? Nous pourrions illustrer

²⁸ Voir le site du Newton Project. Accès : <http://www.newtonproject.sussex.ac.uk/prism.php?id=1>.

des types similaires des conceptualisations, les uns simples, les autres complexes, dans une multiplicité de projets en HN ; que ceux-ci soient de la numérisation, de l'archivage, de l'annotation, de l'analyse, de l'indexation, de la diffusion, etc.

La modélisation conceptuelle n'est cependant pas une tâche facile, car les objets et domaines des HN sont complexes. Ce sont des objets sémiotiques sur lesquels plusieurs points de vue, interprétations, explications peuvent être appliquées. Cette sémiotité pose des problèmes particuliers à la modélisation conceptuelle en lien avec des modélisations formelle et informatique. Tout au long d'un projet de recherche en HN, la modélisation conceptuelle se trouve partout, si bien qu'elle devient transparente. Plusieurs chercheurs penseront peut-être qu'elle ne joue aucun rôle dans la recherche. Or, elle soutient la théorie en ce qu'elle interagit avec les modèles formel et informatique, et elle est essentielle à la compréhension.

Conclusion

Nous avons exploré une hypothèse épistémologique pour mieux comprendre la nature de la démarche de recherche des humanités numériques. Nous nous sommes inspiré de la vision pragmatique de la science pour laquelle les théories scientifiques sont essentiellement constituées de divers types de modèles dont le rôle est de décrire, expliquer et comprendre des réalités, que celles-ci soient naturelles ou humaines. Cette vision de la science a permis d'interroger la possibilité d'explorer la nature épistémique des recherches entreprises dans les humanités numériques.

Tout comme les autres sciences, les HN utilisent plusieurs types de modèles. Nous en avons distingué trois. Un premier contient des modèles formels. Un de ses sous-modèles est de type mathématique. Il identifie des entités et des relations structurantes dans les objets d'étude et s'exprime dans des langages mathématiques spécifiques. Un autre sous-modèle est computationnel. Celui-ci retient des modèles mathématiques les fonctions structurantes qui sont computables. Dans certains cas, si ces fonctions ne sont pas computables, on pourra les remplacer par des heuristiques. Une deuxième classe de modèle est de nature matérielle qui, dans le HN, est surtout de nature informatique. Celle-ci renvoie à une machine électronique, c'est-à-dire un ordinateur, qui peut, causalement, effectuer la computation des fonctions calculables ou des algorithmes retenus dans le modèle formel et exprimés par des algorithmes. Cependant, si la computation concrète est opaque à la cognition, des *simulations* informatiques pourront rendre plus visibles certaines relations de causalité ou mécanismes, permettant ainsi une meilleure compréhension de la réalité étudiée. Enfin, une troisième classe de modèles est de type conceptuel. Ces derniers modèles, exprimés dans un langage naturel, permettent diverses formes de conceptualisation de l'objet de recherche et, à leur tour, ils assistent une compréhension plus limpide de la complexité de la recherche et de ses explications. C'est par ce langage naturel que le modèle conceptuel peut être communiqué, échangé et discuté dans des communautés épistémiques.

Enfin, parce que les divers types de modèles opèrent ensemble et se complètent mutuellement, il arrive que les chercheurs commencent leur recherche par l'une ou l'autre de ces modélisations. Aucun modèle n'est prioritaire comme point de départ d'une recherche. Ainsi, dans certains cas, la recherche pourrait-elle commencer par des modèles formels. Certaines expressions formelles ont été dérivées d'autres expressions formelles et serviront de point de départ pour expliquer des régularités dans l'objet d'étude. Tel est par exemple le cas de la loi de Zipf, dérivée des équations de la théorie de l'information. Dans d'autres situations, la recherche part de données engendrées par des cas sémiotiques. Elle tente alors d'approximer soit par des extrapolations, soit par des algorithmes d'apprentissage, la fonction et son expression formelle : l'équation. C'est souvent le cas lorsque la recherche est placée devant des données massives dont on cherche à extraire des régularités structurantes inconnues.

D'autres recherches partent plutôt de modèles informatiques, c'est-à-dire d'outils existants. Elles les appliquent à des données produites dans les humanités. Par exemple, dans l'étude de grands corpus textuels. Enfin, d'autres recherches partent de modèles conceptuels. Souvent confrontées à des objets sémiotiques d'une grande complexité, avant de passer à des modèles formels ou informatiques, ces recherches visent, dès le départ, à préciser les concepts et les opérations présentes dans l'objet d'étude. Ce n'est que lorsqu'une conceptualisation est stabilisée et acceptée que la recherche explore des modèles formels ou un traitement informatique. Notre propre recherche sur l'analyse conceptuelle assistée par ordinateur (Meunier, Forest, 2009 ; Chartrand et *al.*, 2016) a exigé de nombreuses clarifications conceptuelles pour comprendre plus clairement les principes d'un tel type d'analyse. Ce n'est qu'après cette conceptualisation que des modèles formels et informatiques ont pu être choisis.

Dans une telle perspective, les projets en HN en appellent à plusieurs modèles qui, dans leur application, s'influencent les uns les autres. Et, souvent, leur utilisation est itérative. Les modèles formels, matériels et conceptuels s'influencent mutuellement et imposeront des révisions. Les HN doivent appliquer des principes de génie logiciel où les constantes reprises, itérations, ajustements des modèles sont inhérents à la démarche. Pas plus que les projets typiquement informatiques, les projets HN ne se terminent au premier tour de piste. Et, dans cette démarche, l'ordinateur n'est pas une solution miracle. Il n'est qu'un outil dans la dynamique interprétative qu'est la pratique cognitive scientifique.

Enfin, il faut voir que si les HN incluent dans leur recherche de la modélisation formelle matérielle et conceptuelle, cela les contraint radicalement et délimite les méthodes. Et si le domaine des humanités présentait des relations et des fonctions non computables, l'appel à l'ordinateur serait inutile à moins de lui donner des « oracles ». Dans ce cas, l'ordinateur ne deviendrait qu'un assistant dans l'exploration heuristique des humanités. Et les oracles devront alors laisser leur place à Hermès.

Références

- Armatte M., Dahan Dalmedico A., 2004, « Modèles et modélisations, 1950-2000 : nouvelles pratiques, nouveaux enjeux », *Revue d'histoire des sciences*, 57 (2), pp. 243-303.
- Bachelard S., 1979, « Quelques aspects historiques des notions de modèle et de justification des modèles », pp. 3-19, in : Delattre P., Thellier M., dirs, *Élaboration et justification des modèles*, Paris, Maloine.
- Băetu T. M., 2013, « When Is a Mechanistic Explanation Satisfactory? Reductionism and Antireductionism in the Context of Mechanistic Explanations », *Romanian Studies in Philosophy of Science*, 3, pp. 255-268.
- Bateson G., 1972, *Vers une écologie de l'esprit*, trad. de l'américain par F. Drosso, L. Lot et al., Paris, Éd. Le Seuil, 1977-1980.
- Bechtel W., 2008, *Mental Mechanisms. Philosophical Perspectives on Cognitive Neuroscience*, Londres, Routledge.
- Goodfellow I., Bengio Y., Courville A., 2015, *Deep Learning*, MIT Press.
- Bonaccorsi J., Carayol V., Domenget J.-C., coords, 2016, « Humanités Numériques et Sciences de l'Information et de la communication », *Revue française des sciences de l'information et de la communication*, 8. Accès : <https://rfsic.revues.org/1778>.
- Brown H. I., 2007, *Conceptual Systems*, Londres, Routledge.
- Burch T., K., 1999, « Computer Modelling of Theory: Explanation for the 21st Century », *PSC Discussion Papers Series*, 13 (4), art. 1. Accès : <http://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=1108&context=pscpapers>.
- Busa R. A., 2008, « Foreword: Perspectives on the Digital Humanities », in: Schreiban S., Siemens R., Unsworth J., dirs, *A Companion to Digital Humanities*, Hoboken, Wiley-Blackwell. Accès : <http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-1-2>.
- Cartwright N., 1983, *How the Laws of Physics Lie*, Oxford, Oxford University Press.
- Chaitin G. J., 1998, *The Limits of Mathematics*, Singapour, Springer.
- Chaitin G. J., Doria F. A., Costa N. C. A. da, 2012, *Gödel's Way: Exploits into an Undecidable World*, Boca Raton, CRC Press.
- Chartier J.-F., Meunier J.-G., 2011, « Text Mining Methods for Social Representation Analysis in Large Corpora », *Papers on Social Representations*, 20, pp. 37.1-37.47.
- Chartrand L., Meunier J.-G., Pulizzotto D., López González J., Chartier J.-F., Tan Le N., Lareau F., Trujillo Amaya J., 2016, « CoFiH: A heuristic for concept discovery in computer-assisted conceptual analysis », in: Mayaffre D., Poudat C., Vanni L., Magri V., Follette P., dirs, 2016, *JADT 2016: Proceedings of the 13th International Conference on Statistical Analysis of Textual Data*. Accès : <https://jad2016.sciencesconf.org/82933/document>.
- Chomsky N., Schützenberger M. P., 1963, « The Algebraic Theory of Context Free Languages », pp. 118-161, in: Braffort, Hirschberg D., eds, *Computer Programming and Formal Systems*, Amsterdam, North Holland.
- Church A., 1937, « Review: On Computable Numbers, with an Application to the Entscheidungsproblem by A. M. Turing », *The Journal of Symbolic Logic*, 2 (1), pp. 42-43.

- Craver C. F., 2006, « When mechanistic models explain », *Synthese*, 153 (3), pp. 355-376.
- Craver C. F., Tabery J., 2015, « Mechanisms in Science », in: Zalta E. N., *Stanford Encyclopedia of Philosophy*, 2016. Accès : <https://plato.stanford.edu/entries/science-mechanisms/>.
- Curry H. B., Feys R., 1958, *Combinatory Logic*, vol. I, Amsterdam, North-Holland.
- Dacos M., Mounier P., 2014, *Humanités numériques. État des lieux et positionnement de la recherche française dans le contexte international*, Paris, Institut français/ministère des Affaires étrangères pour l'action culturelle. Accès : http://www.institutfrancais.com/sites/default/files/if_humanites-numeriques.pdf.
- Dennett D. C., 1978, *Brains Storms. Philosophical Essays on Mind and Psychology*, Cambridge, MIT Press.
- Duhem P., 1906, *La Théorie physique. Son objet et sa structure*, Paris, Chevalier/M. Rivière.
- Fraassen B. C. van, 1970, « On the Extension of Beth's Semantics of Physical Theories », *Philosophy of Science*, 37 (3), pp. 325-339.
- Frigg R., Hartmann S., 2006, « Models in Science », in: Zalta E. N., ed., *Stanford Encyclopedia of Philosophy*, 2012. Accès : <https://plato.stanford.edu/entries/models-science/>.
- Gärdenfors P., 2000, *Conceptual Spaces. The Geometry of Thought*, Cambridge, MIT Press.
- Giere R. N., 1999, « Using Models to Represent Reality », pp. 41-57, in: Magnani L., Nersessian L., Thagard P., eds, *Model-Based Reasoning in Scientific Discovery*, New York, Plenum.
- Graziano M. S. A., 2013, *Consciousness and the Social Brain*, Oxford, Oxford University Press.
- Green S., 2013, « When One Model is not Enough: Combining Epistemic Tools in Systems Biology », *Studies in History and Philosophy of Science Part C : Studies in History and Philosophy of Biological and Biomedical Sciences*, 44 (2), 170-180.
- Hesse M. B., 1963, *Models and Analogies in Science*, Londres/New York, Sheed/Ward.
- Jodelet D., dir., 2003, *Les Représentations sociales*, Paris, Presses universitaires de France.
- Johnson-Laird P. N., 1983, *Mental Models. Towards a Cognitive Science of Language, Inference, and Consciousness*, Cambridge, Harvard University Press, 1986.
- Knuuttila T., 2011, « Modelling and Representing: An Artefactual Approach to Model-Based Representation », *Studies in History and Philosophy of Science*, 42 (2), pp. 262-271.
- Kuhn T. S. 1962. *The Structure of Scientific Revolutions*, Chicago, University of Chicago Press.
- Le Moigne J.-L., 1989, « Qu'est-ce qu'un modèle ? », *Confrontations psychiatriques*, numéro spécial, pp. 11-30. Accès : <http://archive.mcxapc.org/docs/ateliers/lemoign2.pdf>.
- Leonelli S., 2007, « What is in a Model? Combining Theoretical and Material Models to develop Intelligible Theories », pp. 15-36, in: Laubichler M. D., Müller G. B., eds, *Modeling Biology. Structures, Behaviors, Evolution*, Cambridge, MIT Press.
- Leplin J., ed., 1984, *Scientific Realism*, Berkeley, University of California Press.
- Markov A., 1960, « The Theory of Algorithms », *American Mathematical Society Translations*, series 2, 15, pp. 1-14.
- Marr D., 1982, *Vision: Computational Investigation into the Human Representation and Processing of Visual Information*, San Francisco, W. H. Freeman.

- McCarty W., 2004, « Modeling: A Study in Words and Meanings », in: Schreibman S., Siemens R., Unsworth J, eds, *A Companion to Digital Humanities*, Hoboken, Blackwell. Accès : <http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-3-7>.
- Meunier J.-G., 2014, « Humanités numériques ou computationnelles. Enjeux herméneutiques », *Sens Public*, 12. Accès : <http://sens-public.org/spip.php?article1121>.
- Meunier J.-G., à paraître, « Vers une sémiotique computationnelle », *Applied Semiotics/ Sémiotique appliquée*.
- Meunier J.-G., Forest D., 2009, « Annotation, lecture et analyse conceptuelle assistée par ordinateur : premières expériences », pp. 211-230, in : Le Priol F., Desclés J.-P., dirs, *Annotations automatiques et recherche d'information*, Cachan/Paris, Hermes/Lavoisier.
- Mitchell T. M., 1997, *Machine Learning*, New York, McGraw-Hill.
- Morgan M. S., Morrison M., eds, 1999, *Models as mediators. Perspectives on natural and social science*, Cambridge, Cambridge University Press.
- Neumann J. von, 1945, *First Draft of a Report on the EDVAC*. Accès : <http://www.virtualtravelog.net/wp/wp-content/media/2003-08-TheFirstDraft.pdf>.
- Newell A., 1980, « Physical Symbol systems' », *Cognitive science*, 4 :2, pp. 135-183.
- Newell A., 1983, « Intellectual Issues in the History of artificial Intelligence », in: Machlup F., Manfield U., eds, *The Study of Information: Interdisciplinary Messages*, New York, J. Wiley.
- Papineau D., 2010, « Realism, Ramsey Sentences and the Pessimistic Meta-Induction », *Studies in History and Philosophy of Science*, 41, pp. 375-385.
- Polanyi M., 1967, *The Tacit Dimension*, New York, Anchor.
- Post E.L., 1936, « Finite Combinatory Processes Formulation », *Symbolic Logic*, 1 (3), pp. 103-105.
- Pylyshyn Z.W., 1984, *Computation and Cognition. Toward a Foundation for Cognitive Science*, Cambridge, MIT Press.
- Rheinberger H.-J., 1997, *Towards a History of Epistemic Things. Synthesizing Proteins in the Test Tube*, Stanford, Stanford University Press.
- Rogers H., 1967, *Theory of Recursive Functions and Effective Computability*, New York, McGraw-Hill, 1987.
- Schmidhuber J., 2015, « Deep Learning in Neural Networks: An Overview », *Neural Networks*, 61, pp. 85-117.
- Soare R. I., 2009. « Turing Oracle Machines, Online Computing, and Three Displacements in Computability Theory », *Annals of Pure and Applied Logic*, 160 (3), pp. 368-399.
- Tarski A., 1933, « The Concept of Truth in Formalized Languages », trad. de l'allemand par J. H. Woodger, pp. 152-278, in: Corcoran J., ed., *Logic, Semantics, Metamathematics*, Cambridge, Hackett, 1983.
- Turing A.M., 1937, « On Computable Numbers, with an Application to the Entscheidungsproblem », in: *Proceedings of the London Mathematical Society*, s2-42, pp. 230-265.
- Turing A. M., 1939, « Systems of Logic Based on Ordinals », *Proceedings of the London Mathematical Society*, s2-45, pp. 161-228.
- Valéry P., 1957-1961, *Cahiers*, 2 volumes, Paris, Gallimard, 1977.

- Weisberg M., 2007, « Who is a Modeler? », *British Journal for Philosophy of Science*, 58 (2), pp. 207-233.
- Weisberg M., 2015, *Simulation and Similarity. Using Models to Understand the World* Oxford, Oxford University Press.
- Winther R. G., 2015, « The Structure of Scientific Theories », in: Zalta E. N., ed., *Stanford Encyclopedia of Philosophy*, 2016. Accès : <https://plato.stanford.edu/archives/win2016/entries/structure-scientific-theories/>.

Sites internet

Tapor, *Discover research tools for studying texts*. Accès : <http://tapor.ca/home>.

The Newton Project. Accès : <http://www.newtonproject.sussex.ac.uk/prism.php?id=1>.