



HAL
open science

Intelligence artificielle et épistémologie. Allers-retours indispensables

Jean-Gabriel Ganascia

► **To cite this version:**

Jean-Gabriel Ganascia. Intelligence artificielle et épistémologie. Allers-retours indispensables. JECIS, Jean-Sébastien Vayre; Gérald Gaglio; Manuel Boutet; Lise Arena, Jun 2022, Nice, France. hal-03760357

HAL Id: hal-03760357

<https://hal.science/hal-03760357>

Submitted on 30 Aug 2022

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.

Intelligence artificielle et

épistémologie

Allers-retours indispensables

Jean-Gabriel Ganascia

Résumé : Ce chapitre montre que les enjeux épistémologiques mutuels de l'intelligence artificielle (IA) et des sciences de l'homme et de la société (SHS) sont multiples. Il en aborde précisément trois. Le premier tient à l'épistémologie de l'IA qui, en tant que discipline scientifique, mérite qu'un regard philosophique et historique en dégage les fondements. Il s'agit de préciser la nature de cette discipline qui ne se réduit pas à une technique et qui, en tant que science, n'est ni une science de la nature à proprement parler, ni vraiment une science de la culture ; de plus, tout en étant à l'origine fondée par des mathématiciens, elle recourt aux SHS. Le deuxième enjeu est relatif aux utilisations de l'IA, de l'apprentissage machine et du traitement des données dans différentes disciplines des SHS et aux changements majeurs que ces utilisations induisent dans ces disciplines en automatisant des tâches fastidieuses. Enfin, le troisième enjeu vient des bouleversements apportés par l'IA et le numérique aux SHS où ils permettent non seulement d'automatiser certaines tâches, mais aussi de concevoir de nouveaux opérateurs d'interprétation et des procédures de preuve. L'apport de l'IA n'est alors pas uniquement pratique ; il introduit dans ces sciences ce que Gaston Bachelard appelle une rupture épistémologique, c'est-à-dire une dissociation entre l'évidence première de l'observation et les faits issus de l'expérimentation.

Mots-clés : épistémologie, humanités numériques, intelligence artificielle, sciences de la culture, sciences de la nature.

Intelligence artificielle et épistémologie

Étude des connaissances scientifiques, l'épistémologie aborde les disciplines sous deux angles au moins, un angle descriptif qui expose les travaux des chercheurs, leurs découvertes et leurs controverses, et un angle prescriptif qui précise les normes auxquelles les scientifiques doivent soumettre leurs investigations pour s'assurer de la rigueur des résultats obtenus. L'intelligence artificielle (IA), en tant que discipline scientifique doit, elle aussi, être examinées sous ces deux angles, l'angle historique et descriptif, qui rend compte de sa constitution et de ses évolutions, d'un côté, l'angle normatif, qui expose les fondements de cette discipline et la nature des connaissances produites, de l'autre. À l'évidence donc, le regard de l'historien et celui du philosophe aideront à mieux comprendre le statut propre de l'IA qui, comme nous tâcherons de le montrer, ne se résume pas à un ensemble de technologies, mais se définit d'abord comme une science particulière.

À ces deux premières dimensions de l'épistémologie de toute discipline scientifique, il convient, lorsqu'on aborde l'IA, d'en ajouter une troisième qui tient à la place particulière qu'elle prend dans l'activité scientifique contemporaine. Comme nous allons le voir, la numérisation des données et leur traitement par des algorithmes d'apprentissage machine, changent la nature de l'enquête scientifique, ce qui mérite une réflexion d'ordre épistémologique. Et, ce qui vaut pour l'activité scientifique en général, vaut encore, mais de façon différente, pour les sciences de l'homme et de la société (SHS). Au-delà d'être un simple outil, l'IA transforme les disciplines. Son apport ne se réduit pas simplement à les équiper. Il a une portée bien plus fondamentale ; l'IA introduit dans tous les domaines, et tout particulièrement dans les SHS, ce que Gaston Bachelard appelle une rupture épistémologique, c'est-à-dire une dissociation entre l'évidence première de l'observation et les faits scientifiques issus de l'expérimentation. Ainsi, dans le champ des humanités numériques littéraires, l'IA permet de tester des champs d'hypothèses sur de nombreux textes, sans se laisser abuser par une ou deux observations judicieuses, comme c'était souvent le cas dans le passé. C'est ce que nous tâcherons de mettre en évidence dans cet article en distinguant deux aspects, l'un pratique, qui souligne les commodités, l'autre plus fondamental, qui tentera d'approcher les évolutions radicales dans le mode d'approche scientifique des SHS.

Après avoir fait un rappel de l'histoire de l'IA, nous aborderons successivement une réflexion sur l'épistémologie de l'IA, puis sur l'apport pratique des SHS dans l'IA avant d'examiner l'apport du numérique dans les SHS et de nous demander comment l'épistémologie des SHS est susceptible d'être modifiée par l'IA.

Préhistoire et histoire de l'IA

Projets antérieurs de simulation de la pensée

Les projets de mécanisation des lois de la pensée et d'automatisation du raisonnement sont anciens. Née dans l'antiquité, la logique visait à déterminer les lois de la pensée juste ; elle caractérisait, au moyen de règles mécaniques, les raisonnements valides comme étant des enchaînements d'inférences — c'est-à-dire de manipulations formelles d'expressions — conformes à des figures élémentaires répertoriées elles-mêmes comme étant valides. En l'occurrence, on déterminait l'ensemble des inférences élémentaires à partir de syllogismes « règlementaires », c'est-à-dire de figures faisant passer de deux propositions à une troisième.

Plus tard, dans la seconde moitié du XVII^e siècle, avec Leibniz, apparut la volonté de mathématiser les lois de la logique pour rendre compte de la validité d'un raisonnement par un calcul, sans avoir à mémoriser l'ensemble des syllogismes élémentaires valides. Immédiatement après, vint le souhait d'automatiser ce calcul sur une machine. De ce fait, on peut dire de Leibniz qu'il est un précurseur de l'intelligence artificielle ; il a d'ailleurs essayé de dessiner les plans d'une machine capable de raisonner d'elle-même.

Ce projet a été repris en XIX^e siècle par George Boole qui créa une algèbre binaire pour rendre compte des lois de la logique, puis par William Stanley Jevons qui construisit une machine, le « piano logique », à même de déduire mécaniquement les conséquences de prémisses logiques en se fondant sur les travaux de George Boole.

Plusieurs autres tentatives d'automatisation du raisonnement virent le jour au début du XX^e siècle. Songeons par exemple aux machines mécaniques jouant aux échecs de Torres y Quevedos qui furent construites l'une en 1912, l'autre en 1920. On doit aussi mentionner le mouvement cybernétique avec entre autres Walter Pitts, Claude Shannon et Norbert Wiener, car il fut aussi à l'origine de plusieurs tentatives de reproduction de la pensée sur des calculateurs électroniques.

Enfin, il ne faut pas oublier Alan Turing qui se demanda, dans son fameux article *Computing Machinery and Intelligence* publié en 1950, ce que signifie penser pour une machine et comme s'y prendre pour fabriquer une telle machine pensante.

Naissance de l'IA

Cependant, même si l'idée que l'on puisse construire une machine reproduisant la pensée existait depuis longtemps, le terme « intelligence artificielle » n'apparut pour la première fois qu'en 1955 dans un projet d'école d'été soumis par quatre chercheurs, John McCarthy, Marvin Minsky, Nathanael Rochester et Claude Shannon, à la fondation Rockefeller pour obtenir une subvention afin d'organiser une école d'été au « Dartmouth College », dans le New Hampshire aux États-Unis en 1956.

Pour les promoteurs de cette école d'été, l'intelligence artificielle désignait une discipline scientifique qui visait à étudier l'intelligence avec les ordinateurs. Cette nouvelle science reposait sur la *conjecture* selon laquelle toutes les facultés cognitives, en particulier le raisonnement, le calcul, la perception, la mémorisation, voire la découverte scientifique ou la créativité artistique, pourraient être décrites avec une précision telle qu'il devrait être possible de les reproduire à l'aide d'un ordinateur. Insistons sur l'importance épistémologique de cette conjecture : elle dessine un horizon de tâches à accomplir, à l'instar du postulat de Galilée selon lequel *la nature serait écrite en langue mathématique*. Soulignons enfin que cela ne signifie pas que l'on cherche à fabriquer une machine intelligente qui serait un double de nous-même.

Soixante-cinq ans plus tard, et en dépit des progrès époustouflants enregistrés, l'étude de l'intelligence artificielle repose toujours sur la même conjecture que rien, jusqu'à présent, n'a permis ni de démentir, ni de démontrer irréfutablement.

Différentes phases de l'histoire de l'IA

Née en 1956, l'intelligence artificielle a connu de nombreuses évolutions au cours de sa courte existence. On peut les résumer en six étapes.

Le temps des prophètes

Quelques réalisations, en particulier le LOGIC THEORIST, ce programme informatique réalisé en 1957 qui démontrait automatiquement des théorèmes de logique, enthousiasmèrent. Dans l'euphorie qui s'ensuivit, les chercheurs se sont laissés aller à des déclarations un peu inconsidérées qu'on leur a beaucoup reprochées par la suite.

L'hiver de l'IA

Au milieu des années 1960, les progrès tardèrent ; l'intelligence artificielle eut alors mauvaise presse, et c'est ainsi que commencèrent quelques années sombres. C'est ce que l'on a appelé

l'hiver de l'IA. Cela s'est traduit par une diminution des financements des laboratoires de recherche en IA aux États-Unis. En Europe, en particulier en France et en Allemagne, il y avait de toutes les façons très peu d'équipes travaillant sur ce domaine, sauf au Royaume-Uni. Les travaux ne s'interrompirent pas pour autant, en particulier dans le domaine de la démonstration automatique de théorèmes. Ainsi, à cette époque, en France, Alain Colmerauer conçut le langage PROLOG. Mais, il disposait de très peu de financement, car l'IA, comme l'informatique en général, était vue surtout comme une discipline spéculative et théorique, au même titre que les mathématiques.

L'intelligence artificielle sémantique

Dans le milieu des années soixante-dix, aux États-Unis, on axa les recherches dans de nouvelles directions en s'intéressant à la mémoire et aux mécanismes de compréhension, que l'on chercha à simuler sur un ordinateur. On s'interrogea aussi sur le rôle de la connaissance dans le raisonnement. C'est ce qui donna naissance aux techniques de représentation des connaissances et aux systèmes dits experts, parce qu'ils recourraient au savoir d'hommes de métiers pour reproduire leurs raisonnements. Ces derniers suscitèrent d'énormes espoirs au début des années 1980. Ce courant donna naissance à ce que l'on appela l'intelligence artificielle sémantique qui se développa, comme nous allons le voir, en lien avec les SHS.

Néo-connexionnisme et apprentissage machine

Parallèlement à l'essor de l'intelligence artificielle au début des années 1980, les techniques issues de la cybernétique et du connexionnisme se perfectionnèrent, s'affranchirent de leurs limitations initiales et firent l'objet de multiples formalisations mathématiques. Ce fut l'origine des algorithmes d'apprentissage dits de rétropropagation de gradient pour les réseaux de neurones formels.

De l'intelligence artificielle à l'« informatique animiste »...

À partir de la fin des années 1990, on coupla l'intelligence artificielle à la robotique et aux interfaces homme-machine, de façon à produire des agents intelligents qui suggèrent la présence d'un autre. Cette tendance de l'IA peut éventuellement se caractériser comme une forme d'animisme en cela qu'elle s'emploie à susciter la projection d'un souffle de vie sur les objets quotidiens de notre environnement.

Renaissance de l'intelligence artificielle

Depuis environ 2010, la puissance des machines permet d'exploiter des grandes masses de données (ce que l'on appelle couramment les *Big Data*) avec des techniques d'apprentissage machine.

Place des SHS dans l'IA

Les pionniers de l'IA étaient plutôt mathématiciens ou ingénieurs de formation. Il y eut toutefois dès le début, des spécialistes des SHS qui contribuèrent à l'avancement de la discipline. Ainsi, lors de l'école de Dartmouth college, il y avait parmi les vingt participants, un psychiatre, Ross Ashby, un neurophysiologiste, Warren McCulloch, un chercheur en psychologie cognitive, Allan Newell, et un docteur en sciences politiques qui devint, plus tard, prix Nobel d'économie, Herbert Simon. Et, au cours de l'histoire de l'IA, les SHS prirent de temps à autres plus d'importance. C'est ce que nous allons expliquer dans la suite de cette section.

Approche logico-mathématique

Les premiers travaux d'IA, dans les années cinquante et soixante, reposaient surtout sur des modélisations mathématiques, en particulier sur les statistiques et sur la logique. Ce furent le cas des démonstrateurs automatiques de théorèmes, par exemple du LOGIC THEORIST d'Herbert Simon et d'Allan Newell, de la résolution de problème et des premières tentatives d'apprentissage machine d'Arthur Samuel, en particulier de l'apprentissage par renforcement. Cela donna naissance à une science des modèles.

En parallèle, un courant recourrait à la psychologie comportementale afin d'évaluer la vraisemblance des modélisations cognitives du raisonnement ou de l'apprentissage. Ce fût par exemple, l'origine des premiers travaux d'apprentissage non supervisé ou d'apprentissage par l'action, avec le « morcelage » (*chunking* en anglais).

Approches sémantiques

À partir des années soixante-dix, un nouveau courant de l'IA s'inspira de la psychologie cognitive, en particulier de la théorie des schémas de Charles Bartlett et de la théorie du prototype d'Eleanor Rosch, ainsi que de la linguistique, avec les grammaires transformationnelles issues des théories de Chomsky d'un côté et, surtout, des grammaires de cas de Fillmore et de la grammaire sémantique de Montague de l'autre, afin de mieux comprendre les facultés cognitives humaines avant de les modéliser. Les techniques de représentation des connaissances, en particulier les réseaux sémantiques, et les systèmes à base de connaissance, dits aussi systèmes experts, sont directement issus de ces travaux.

Il en est résulté une tension entre deux vues de l'intelligence artificielle, l'une qui se centrait plus sur les propriétés logico-mathématiques requises pour qu'une simulation de raisonnements sur machine soit possible, l'autre sur l'étude des processus psychiques et linguistiques que l'on souhaite modéliser. Cette tension a été résolue au début des années quatre-vingt avec la

formalisation logique des techniques de représentation des connaissances, en particulier avec les logiques de description qui sont aujourd'hui au fondement de ce que l'on appelle les ontologies formelles.

Théories de l'apprentissage et apprentissage profond

À partir des années quatre-vingt et de la mise en œuvre de nombreux modèles d'apprentissage supervisé, il y eut des tentatives de théorisation mathématique de l'apprentissage machine avec, en particulier les travaux de Leslie Valiant sur la théorie de l'apprenable et de Vladimir Vapnik sur l'apprentissage statistique. Ces approches furent à l'origine de nouvelles approches, en particulier les méthodes d'ensemble (« bagging » et « boosting ») et les machines à vecteurs de supports (SVM).

Il en est résulté un grand nombre d'applications dans les SHS, en particulier en linguistique où l'on a pu exploiter automatiquement de très grands corpus pour en tirer des connaissances.

Masse de données

Depuis le début du XXI^e siècle, les opérateurs de l'internet utilisèrent les techniques d'apprentissage machine pour traiter de très grandes masses de données qui se comptent en giga- (10^9), en téra- (10^{12}), voire en péta-octets (10^{15}). On s'est ensuite rendu compte, à partir de 2010, que les approches pragmatiques qui faisaient appel à des réseaux de neurones formels organisés en multiples couches donnaient des résultats statistiques bien supérieurs aux modèles antérieurs, sans disposer de théorie mathématique qui explique ces résultats, sans reproduire non plus ce que l'on connaît du fonctionnement du cerveau. Rien ne dit toutefois que l'on ne disposera pas d'une telle théorie dans le futur. Là encore, les SHS ont tiré un très grand parti de l'IA.

Utilisation de l'IA dans les SHS

Premières utilisations de l'informatique dans les SHS

Le concept d'ordinateur, c'est-à-dire de machine à exécuter des algorithmes, a été conçu au tout début du XIX^e siècle par un économiste anglais, Charles Babbage. Son assistante, Ada Byron Lovelace, la fille du grand poète George Gordon Byron, eut très précocement l'intuition que ces machines ne serviraient pas uniquement à faire des calculs mathématiques, mais qu'elles traiteraient aussi des textes et de la musique tant pour assister les créateurs que pour mieux comprendre l'activité créatrice. À la

lumière de l'intuition géniale d'Ada Byron, on conçoit rétrospectivement qu'on a pressenti très tôt l'utilité de l'informatique dans les disciplines d'érudition et plus généralement dans les SHS.

D'ailleurs, dès que sont apparus les premiers ordinateurs électroniques, on y a recouru pour assister des lettrés dans leurs travaux. On cite souvent le jésuite italien Roberto Busa, qui, en 1949, à peine trois ans après la réalisation du premier ordinateur électronique, amorce un patient travail de transcription informatique de la *Somme théologique* de Thomas d'Aquin. Depuis, les utilisations des ordinateurs dans les humanités, puis dans les SHS n'ont pas cessé, donnant naissance aux sciences cognitives dans les années soixante-dix puis à la communauté informatique pour les sciences humaines dans de multiples disciplines comme l'archéologie avec, en France, les travaux précurseurs de Jean-Claude Gardin, à l'EHESS, dans le cadre du LISH, ou en histoire, avec les approches pionnières de Gian Piero Zatti, ou encore en littérature, dès les années soixante, avec des outils de stylistique automatique fondés sur l'analyse de données, ainsi que de nouveaux outils d'interprétation conçus pour faciliter le travail dans les disciplines d'érudition, comme le projet TUSTEP [*Tübingen System of Text Processing Programs*] lancé en 1966.

Recourt à l'IA

Quelques années plus tard, la linguistique de corpus et la sémantique se sont considérablement développées grâce aux analyses d'IA sur la représentation des connaissances et sur l'apprentissage automatique. De même, de nombreux développements ont ouvert des perspectives sur les études littéraires. Au début du millénaire, sous la pression d'éditeurs (et non de scientifiques) le terme anglais *Humanities Computing* qui caractérisait les recherches à la frontière des humanités au sens anglo-saxon, c'est-à-dire des sciences de la culture, se transforma en *Digital Humanities* (« humanités numériques »). Ce changement de dénomination signifiait que les humanités n'étaient plus simplement instrumentées par l'utilisation d'ordinateurs, mais qu'elles se transformaient en devenant numérique. On a parlé de « tournant calculatoire » pour caractériser cette évolution vue comme fondamentale. Le renouvellement qui s'en est suivi a été très fructueux. Dans tous les pays, de nombreuses équipes de recherche travaillant dans le champ des humanités classiques, que ce soit en histoire, en archéologie ou en littérature, ont renouvelé leurs approches en collaboration avec des informaticiens.

Plutôt que de broser un tableau, nécessairement incomplet, de toutes les approches dans ce domaine, je propose de me restreindre

ici à un certain nombre de travaux que j'ai réalisés dans le cadre de mon équipe sur le versant littéraire des humanités numériques.

Réécritures, citations, réemplois

Beaucoup de théories de la littérature insistent sur les reprises textuelles, qu'il s'agisse des réécritures d'un même texte, par le même auteur qui, pour paraphraser Boileau, « cent fois sur le métier remet son ouvrage », ou des emprunts plus ou moins conscients, soit littéraires, soit conceptuels à d'autres auteurs. Avec les années, nous avons développé dans mon équipe trois outils destinés à repérer différents types de reprises.

Le logiciel MEDITE (Machine EDITE) a été conçu dans le cadre du projet EDITE (Étude Diachronique et Interprétative de TExtes) en collaboration avec l'ITEM (Institut des Textes et Manuscrits Modernes). Ce laboratoire spécialisé dans la genèse textuelle examine les brouillons et les différents « avant-textes » des grands auteurs, c'est-à-dire les états des œuvres avant publication. À cette fin, MEDITE détecte automatiquement les transformations qui font passer d'une version à une autre, en particulier les *insertions*, les *suppressions*, les *remplacements* et, surtout, les *déplacements*. Dans le passé, de nombreux travaux portèrent sur la comparaison de textes. Cependant, le problème ne reçoit pas de solution optimale : il faut procéder à des compromis et des choix, qu'autorise l'approche heuristique, fondée sur des techniques d'IA, couplés à des principes d'algorithmique des chaînes.

Le logiciel PHŒBUS (Projet d'Hypertexte de l'Œuvre de Balzac utilisant des Similarités) a été conçu dans le cadre d'un projet interdisciplinaire monté par des spécialistes de Balzac et par l'équipe ACASA du LIP6 que je dirige afin de repérer automatiquement des réutilisations et des citations sur de grandes masses de textes. À cette fin, il fait appel à des principes utilisés pour la détection de plagiat qu'il assouplit en utilisant des techniques de traitement automatique du langage naturel recourant à de l'IA, en vue de déceler des reprises approximatives. Ce projet a ensuite été fusionné avec le projet TextAlign de l'université de Chicago, dans le cadre du projet ARTFL.

DeSeRT (Détection Sémantique de Reformulations et de Topiques) est un moteur de recherche sémantique qui découpe les textes en blocs se recouvrant partiellement, puis qui indexe chacun de ces blocs avec les lemmes des mots significatifs (noms, verbes ou adjectifs) présents dans ce bloc. Il repère ensuite des conjonctions récurrentes de termes en ayant éventuellement recours à un thésaurus et à un dictionnaire analogique. Cela permet de repérer sur de grandes masses de textes les arguments qui se répètent et de localiser les blocs où ces arguments apparaissent.

Études stylistiques

Les études littéraires ne se restreignent pas à l'étude des réemplois ou des réécritures, loin de là. À titre d'exemple, de nombreuses approches tentent de caractériser le style littéraire. C'est là encore le cas de plusieurs travaux conduits dans mon équipe ou dans d'autres.

Dans le passé, beaucoup d'études du style individuel des auteurs ont recouru à des statistiques lexicales. Nous avons tenté de développer, dans le cadre de l'équipe ACASA, une stylistique syntaxique, en extrayant automatiquement des motifs syntaxiques récurrents, et une stylistique sémantique, en inventoriant et en caractérisant les figures de comparaison.

Relevant essentiellement de la fouille de données séquentielles, certains travaux portent sur l'extraction et la recherche de motifs syntaxiques à partir de grandes masses de textes. Cela a donné naissance au logiciel EREMOS. Ces extractions automatiques de motifs syntaxiques ont donné prise à des études sur la répartition statistique des motifs à l'intérieur d'œuvres littéraires. À titre d'illustration, une application d'EREMOS conduite en lien étroit avec les équipes de littérature de la Sorbonne, porta sur l'analyse des motifs stylistiques caractéristiques des différents personnages des pièces de théâtre de Molière. Cela a permis d'associer à chaque figure (*Sganarelle*, *Arlequin*, *Don Juan*, etc.) un mode d'expression particulier que l'on est en mesure de caractériser syntaxiquement.

Dans un registre différent, nous avons procédé à l'extraction automatique de figures de comparaison en déterminant, pour chacune, le *comparant*, le *comparé* et le *motif*. Le but est à la fois d'établir automatiquement un dictionnaire des comparaisons, avec leurs sources, et d'évaluer, pour chaque auteur, la richesse des comparaisons, leur nombre et leur répartition.

Transformation des SHS avec l'IA

Au fil des ans, il s'est trouvé que, débordant sa définition originelle, l'IA ne s'est pas limitée à étudier l'intelligence en la décomposant en fonctions cognitives, puis en simulant chacune d'entre elles et en exploitant ces simulations à des fins technologiques. L'IA a aussi joué un rôle clef dans de nombreuses disciplines scientifiques. Ces quelques exemples montrent la variété et la richesse potentielle des apports pratiques de l'IA aux SHS. Dans tous ces cas de figure, nous voyons que les approches de l'IA font référence à des théories préexistantes qu'elle modélise, par exemple les théories de l'intertextualité ou de la génétique textuelle. Ce faisant, et comme nous l'avons vu, l'IA assiste les chercheurs en

SHS en leur fournissant des outils qui automatisent l'exécution de tâches fastidieuses. Au-delà, elle transforme non seulement l'activité scientifique, mais le fondement épistémologique des disciplines elles-mêmes. C'est ce que nous allons essayer de montrer dans cette dernière partie en expliquant pourquoi ces évolutions sont vraiment spécifiques aux SHS.

Sciences de la nature vs sciences de la culture

Pour appréhender cette spécificité des changements épistémologiques des SHS, en les distinguant des transformations advenues dans les sciences dites « dures », nous allons nous appuyer sur l'opposition introduite par des philosophes néo-kantien, en l'occurrence Heinrich Rickert et Ernst Cassirer, au début du XX^e siècle, entre les *sciences de la nature* qui portent sur le monde tel qu'il nous apparaît et les *sciences de la culture* qui étudient les œuvres humaines. Ceux-ci montrent que tant les sciences de la nature que les sciences de la culture sont des sciences empiriques, c'est-à-dire fondées sur des faits d'observation, mais que, des unes aux autres, la logique diffère. Les sciences de la nature visent à construire des lois générales par *induction* à partir d'observations, en oubliant les cas particuliers, tandis que les sciences de la culture se centrent sur les cas particuliers pour leur donner sens en les expliquant.

Transformations des sciences de la nature

L'informatique et l'intelligence artificielle transforment les sciences de la nature en donnant naissance aux *e-sciences*, qui construisent des lois générales à partir de données issues d'observations et qui ont recours à des procédures d'apprentissage machine. Ces dernières font de l'induction au sens logique, c'est-à-dire en passant du particulier au général.

Plus précisément, aujourd'hui, dans de nombreuses disciplines des sciences de la nature, la collation des données se fait automatiquement, sans idée *a priori*. Pour cela on fait appel à des capteurs qui procèdent à des recueils d'observations sans intervention humaine. On exploite ensuite ces données à l'aide d'outils d'IA et d'apprentissage machine, en accomplissant ce que l'on appelle des expériences *in silico*, c'est-à-dire des expériences conduites sur les puces de silicium au moyen de systèmes de traitement de l'information, par opposition aux expériences *in vivo*, qui se font sur le vivant, et *in vitro*, dans des éprouvettes de verre. Ces expériences testent systématiquement de grands nombres d'hypothèses sur les données. Nous assistons là, dans les sciences

de la nature, à une révolution d'ordre épistémologique qui change le statut de l'expérience. En effet, là où l'expérience se construisait à partir d'une théorie initiale qu'il s'agissait de valider ou d'invalider en confrontant les observations recueillies par un dispositif matériel conçu tout exprès avec ce qu'anticipait la théorie initiale, on part désormais, dans les sciences de la nature d'observations massives (par exemple de données issues de capteurs en astrophysique ou du séquençage de protéines ou de génomes en biologie moléculaire) sur lesquelles on procède à des expérimentations virtuelles (expérimentations *in silico*). La notion de preuve, et en particulier de preuve empirique, s'en trouve malmenée...

Transformation des sciences de la culture

À cette évolution radicale des sciences de la nature, fait pendant une transformation majeure des sciences de la culture, à savoir des disciplines qui étudient les œuvres humaines.

Dans ce cas, il ne s'agit plus d'extraire des lois générales par induction à partir de données, mais d'interpréter des cas singuliers, par exemple des œuvres littéraires ou des épisodes historiques, en tirant parti d'une grande multiplicité de données en vue de les comprendre, ou plus exactement à leur donner sens. À cette fin, on adopte une démarche fondée sur ce que les logiciens appellent l'*abduction*, c'est-à-dire sur la recherche d'explications au regard de théories générales. Ainsi, toujours pour la littérature, peut-on tenter de repérer des marqueurs d'influence dans les écrits de grands auteurs. Dès lors, il devient loisible de valider empiriquement certaines hypothèses, ce qui renouvelle les disciplines d'érudition traditionnelles.

Pour aider à cette recherche d'explications, ont été développées et déployées un certain nombre d'outils qui procèdent à de multiples opérations, par exemple à la comparaison d'états de textes ou à la recherche de réemplois, ou encore, en archéologie, à la reconstruction tri-dimensionnelle de poterie ou de bâtiments etc. Ce faisant, ces outils ne se contentent pas d'automatiser des tâches existantes. Ils proposent des opérateurs d'interprétations inédits qui transforment totalement les disciplines des SHS. À titre d'illustration, dans le domaine littéraire, Franco Moretti introduit la notion de lecture distante où il repère des caractéristiques générales sur de grands corpus, telle la longueur des phrases. De même, on peut caractériser les citations ou les emprunts sur de grandes masses de textes, toujours dans le domaine littéraire.

Ces nouveaux opérateurs d'interprétation ont alors un double apport. Certains ont une fonction purement heuristique en suggérant des pistes de recherches nouvelles qui demandent à être approfondies. Ils aident alors à mettre en évidence des

phénomènes jusque-là occultés, en donnant à voir les œuvres humaines sous des jours inédits. Ces pistes de recherche demandent alors à être investiguées de façon plus rigoureuse, avec des méthodologies éprouvées.

D'autres apportent des éléments empiriques de validation ou d'invalidation d'hypothèses de travail, par exemple, dans le champ littéraire, en mettant en évidence certaines influences, ou au contraire en montrant l'absence de références explicites et/ou implicites et de citations. Dans cette dernière éventualité, le fondement scientifique même de certaines disciplines s'en trouve fortement modifié, parce que, comme dans les sciences de la nature, la notion même de preuve évolue avec l'introduction de l'IA dans les sciences de la culture.

Conclusion

Pour conclure rappelons qu'un des pionniers de l'IA, Herbert Simon, a écrit un ouvrage intitulé *The Sciences of the Artificial* où il aborde les approches scientifiques de la modélisation et la fonction des modèles dans les sciences. Cela pourrait conduire à se demander ce qui caractérise l'IA, en tant que science : est-ce une discipline exclusivement théorique, fondée les mathématiques, ou empirique ? Et, dans cette dernière éventualité, s'apparente-t-elle plus aux sciences de la nature ou, pour reprendre la terminologie des philosophes néo-kantien déjà cités comme H. Rickert ou E. Cassirer, aux sciences de la culture, c'est-à-dire aux SHS ? Ce qui nous fait pencher pour cette dernière éventualité, c'est que l'IA porte, en grande partie, sur la modélisation de pratiques individuelles délibérées qui sont le résultat d'activités conscientes et peuvent, en conséquence, être vues comme des œuvres humaines. En tant qu'étude des œuvres humaines, il s'agit donc d'une science de la culture, au sens où le terme a été défini plus haut. Sans doute, les méthodes qu'elle emploie se fondent-elles essentiellement sur des approches mathématiques et statistiques. Mais, en même temps, du point de vue logique, une grande partie de l'activité en IA consiste à rendre compte et à simuler des tâches qui sont le fruit de ces pratiques humaines mentionnées ici et qui, en tant que telles relèvent de la culture. Ainsi, l'étude des relations entre IA et SHS conduit non seulement à montrer l'intérêt historique de l'IA pour les SHS, à mettre en évidence d'utilisation de l'IA par les SHS et les modifications de ces dernières, avec l'IA, ou ce que l'on appelle le « tournant calculatoire » (*computational turn*) de ces dernières, mais aussi la proximité entre l'IA et les SHS.

Bibliographie

- Chris Anderson, “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete”, *Wired Magazine*, 23 juin 2008, <https://pdodds.w3.uvm.edu/files/papers/others/2008/anderson2008a.pdf>
- Barbrook A., Blake N., Howe C., Robinson P. M. W., « *The Philology of the Canterbury Tales* », *Nature*, 394, 1998, p. 839.
- Crasson, A., « Silences, retraits, soupirs : les nervures de l'écriture ». dans *Récit : les cinq états du manuscrit. I et II*, présentation, transcriptions et lectures de Marcel Cohen, Aurèle Crasson, Irène Fenoglio, Paris, Textuel, 2005.
- Frontini, F., Boukhalel, M. A., Ganascia, J.-G., “Mining for Characterizing Patterns in Literature Using Correspondance Analysis. An Experiment on French Novels”, *Digital Humanities Quarterly* Proceedings of the Göttingen Dialogue for Digital Humanities 2015. 11-2 (2017).
- Ganascia, J.-G. & Mainardi, C., “Crossed Semantic Analysis of Literary Texts with DeSeRT”, *Digital Humanities*, Jul 2016, Krakow, Poland.
- Ganascia, J.-G., Glaudes, P. & Del Lungo, A., “Automatic detection of reuses and citations in literary texts”, *Literary and Linguistic Computing*, 29(3) :412–421, 2014.
- Ganascia, J.-G., “The logic of the big data turn in digital humanities”, *Frontiers in Digital Humanities*, 2(7) :1 – 5, 2015.
- Ganascia, J.-G., “In silico' Experiments: Towards a Computerized Epistemology”, in Newsletter on Philosophy and Computers, Piotr Boltuc (ed), American Philosophical Association Newsletters, 07 (2), 11-15, Spr. 2008, ISSN 1067-9464,
- Ganascia, Jean-Gabriel, *Extraction of Recurrent Patterns from Stratified Ordered Trees*, ECML (European Conference on Machine Learning), 2001, p. 167-178.
- Ganascia Jean-Gabriel, « EDITE-MEDITE : un passage des versions aux variantes », dans *Actes du XXIV^e Congrès International de Linguistique et de Philologie Romanes* (Aberystwyth, 1^{er}-6 août 2004), Trotter (David) éd., Tübingen, Niemeyer, 2007.
- Garron I., Minel J.-L., Souchier E., « Citer, indexer ou cartographier ? De la circulation et de la lecture des textes relatifs à une œuvre littéraire sur internet », http://halshs.archivesouvertes.fr/docs/00/08/28/52/PDF/ArticleFinal_IG_JLM_ES.pdf.
- Herbert Simon, (1996). *The Sciences of the Artificial*. MIT Press, Cambridge, Massachusetts, 3rd edition.

- Lebart L., Salem A., *Statistique textuelle*, Paris, Dunod, 1994.
- Newell, A.; Shaw, J. C.; Simon, H. A. (1958). "Elements of a Theory of Human Problem Solving". *Psychological Review*. 65 (3): 151–166
- Newell, A. & Simon, H. A. (1956), "The logic Theory Machine. A Complex Information Processing System", Rand Corporation, P-868,
<http://shelf1.library.cmu.edu/IMLS/MindModels/logictheorymachine.pdf>
- Alan Newell, "The Knowledge Level," *Artificial Intelligence*, 18 (1982), 87-127.
- Ott, W., *Strategies and Tools for Textual Scholarship: the Tübingen System of Text Processing Programs (TUSTEP)*, *Literary and Linguistic Computing* 15, 2000, N° 1, p. 93–108.
- Rosenblueth, A., Wiener, N. & Bigelow, J., "Behavior, purpose and teleology", *Philosophy of Science*, 10, 18–24, (1943).
- Schreibman S., Siemens R. J., Unsworth J. M., *A Companion to Digital Humanities Par*, Wiley-Blackwell, 2004.
- Simon, H. & Newell, A., "Heuristic problem solving: the next advance in operations research", *Operations Research*, 6, 1–10, (1958).
- Turing, A., "Computing machinery and intelligence", *Mind*, 59, 433–460, (1950).
- Vladimir Vapnik, *The Nature of Statistical Learning Theory*, Springer Verlag, 1999
- Zarri, G. P., *Building the Inference Component of an Historical Information Retrieval System*, IJCAI, 1981, p. 401-408.