

## LA PRÉDICTION ALGORITHMIQUE COMME ACTIVITÉ SOCIALE

Tyler Reigeluth

La Découverte | « Réseaux »

2018/5 n° 211 | pages 35 à 67

ISSN 0751-7971

ISBN 9782348040689

Article disponible en ligne à l'adresse :

-----  
<https://www.cairn.info/revue-reseaux-2018-5-page-35.htm>  
-----

Pour citer cet article :

-----  
Tyler Reigeluth, « La prédiction algorithmique comme activité sociale », *Réseaux*  
2018/5 (n° 211), p. 35-67.  
DOI 10.3917/res.211.0035  
-----

Distribution électronique Cairn.info pour La Découverte.

© La Découverte. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

# DOSSIER

---

## Machines prédictives

# LA PRÉDICTION ALGORITHMIQUE COMME ACTIVITÉ SOCIALE

Tyler REIGELUTH

La vie quotidienne abonde de situations où nos comportements sont infléchis plus ou moins insensiblement par les comportements de machines douées de ce qui était jusqu'alors réservé aux êtres vivants : l'apprentissage. Apprendre de nos données pour nous dire quel livre nous aimerions lire, apprendre d'une masse de pixels pour déterminer quel comportement présente une menace à l'intérieur d'un espace surveillé, apprendre des habitudes de vie pour calculer le juste prix d'une assurance santé, etc. Un nombre croissant d'activités sociales est aujourd'hui profondément affecté par l'apprentissage algorithmique, ce sous-domaine de l'intelligence artificielle appelé *machine learning* qui développe des algorithmes capables d'apprendre de leurs expériences afin d'adapter leurs comportements pour résoudre un problème ou améliorer leurs performances sur une tâche. Traditionnellement, un *algorithme est défini comme une suite d'opérations ou d'étapes qui ne peuvent manquer de donner un résultat donné dans un temps donné*. Sous cette acception, l'algorithme n'est pas autre chose qu'un programme dont les opérations sont écrites à l'avance et qui certifie, à partir d'un espace de calcul, un certain déroulement temporel<sup>1</sup>. La fonction même de l'algorithme est de rendre prévisible un ordre d'opérations à partir d'une logique formelle. Mais lorsque nous ne savons pas quelle fonction nous voulons exécuter, il faut l'apprendre à partir des données. C'est cet apprentissage que le *machine learning* automatiserait. Les algorithmes apprenants sont généralement observés et décrits en termes des comportements qu'ils exhibent ; leurs fonctionnements internes sont de plus en plus traités comme une boîte noire, y compris pour les ingénieurs, statisticiens et informaticiens qui participent à leur programmation initiale. Nous assistons dès lors à un déplacement de la fonction (entendu à la fois au sens mathématique et social) algorithmique : d'un algorithme comme

---

1. « Un algorithme n'est pas autre chose en effet que la prescription d'un déroulement temporel à partir d'un code consignée dans l'espace d'un calcul » (Bachimont, 2010, p. 143). Et plus loin : « [...] un programme n'est pas autre chose qu'un dispositif réglant un déroulement dans le temps, le calcul ou l'exécution du programme, à partir d'une structure spécifiée dans l'espace, l'algorithme ou le programme. L'algorithme spécifie que, les conditions initiales étant réunies, le résultat ne peut manquer d'être obtenu, selon une complexité donnée » (*ibid.*, p. 167).

ensemble d'étapes déterminées à exécuter, on est aujourd'hui confronté à la nécessité de considérer un *système* algorithmique aux comportements partiellement imprévisibles. C'est à partir de ce déplacement qu'on peut penser à nouveaux frais la question de l'activité sociale et la normativité de la prédiction algorithmique.

Il n'est plus seulement question, avec le *machine learning*, d'automatiser le *geste* du travail manuel, ou le traitement cognitif de l'*information*, mais d'automatiser le caractère prédictif de tout *comportement*. En tant que régime de production de prédictions, le *machine learning* nous intéresse en ce que son efficacité normative ne se réduit pas à l'efficacité d'une certaine forme discursive, mais s'insère dans une *activité sociale*. L'algorithme ne peut plus être seulement compris, dans un geste classique, comme un discours à partir duquel la norme s'énonce ou auquel elle se rapporte. L'extension de l'apprentissage algorithmique à un nombre croissant d'activités sociales s'accompagne d'une généralisation de la prédiction comme mode de gouvernement des comportements. La plus-value à la fois économique et cognitive de cette automatisation reposerait sur le fait que son fonctionnement ne se limite pas à confirmer ce que l'on sait déjà, mais à produire des effets imprévisibles et à réinventer le sens de l'activité à laquelle elle participe (Mackenzie, 2015). En commentant « l'automatisation » du traitement de données à des fins de prédiction et d'exclusion statistiques sur base de *patterns* en eux-mêmes insignifiants, Antoinette Rouvroy et Thomas Berns (2013) relèvent une disgrâce de la moyenne – en tant qu'opération statistique – propre à la « gouvernementalité algorithmique » contemporaine :

« [C]ette émancipation par rapport à toute forme de moyenne tient notamment au caractère auto-apprenant de ces dispositifs, et peut être considérée comme essentielle à l'action normative contemporaine. C'est de ce point de vue que nous pouvons également dire que la gouvernementalité algorithmique rompt avec l'origine conventionnelle de l'information statistique, telle que décrite par Alain Desrosières [...] : "L'information statistique ne tombe pas du ciel comme pur reflet d'une 'réalité' antérieure à elle. Bien au contraire elle peut être vue comme le couronnement provisoire et fragile d'une série de conventions d'équivalence entre des êtres qu'une multitude de forces désordonnées cherche continuellement à différencier et à disjoindre." » (Rouvroy et Berns, 2013).

En tant que mesure reposant sur une convention (Desrosières, 2014), la moyenne statistique recèle une facticité toujours susceptible d'être sujette à une critique épistémologique et à une réappropriation sociale. L'automatisation

algorithmique, en revanche, serait une modélisation statistique de plus en plus immanente au tissu normatif du monde social. L'écart représentationnel qu'instituait la statistique classique entre le « monde » et le « modèle » pâlit à la lumière d'un réel numérique clos sur lui-même et qui est indifférente à un « dehors » lui échappant et dont l'instauration aurait été, selon Rouvroy et Berns, le geste politique par excellence. Il n'est pas question de statuer sur la justesse de cette hypothèse dans l'absolu, mais il semblerait néanmoins qu'elle soit symptomatique d'un certain aplatissement de la « normativité contemporaine » aux discours sur le « caractère auto-apprenant des dispositifs » algorithmiques. Ainsi, la norme technique se présente d'emblée comme univoque et écrite à l'avance ; elle est un *programme*. Ce faisant, on passe à côté de la technicité elle-même comme source de normativité échappant en partie à l'ordre du discours. On risque alors de tomber dans le piège tendu par l'adéquation préconçue entre norme technique et automatisme selon laquelle suivre une normativité technique reviendrait à automatiser l'activité normative elle-même, à désamorcer toute charge politique à partir de laquelle la différence normative puisse s'éprouver comme un problème réclamant une décision, un partage dont il est possible de contester la légitimité, la pertinence ou la raison d'être. En ce qui nous concerne, reconduire l'adéquation entre automatisme et technique est avant tout une manière de *normaliser* le fonctionnement d'une technique à partir de l'ordre du discours, ou de normes économiques et sociales. L'opération technique est ainsi considérée comme une simple courroie de transmission normative. En revanche, il nous semble que le *machine learning* se présente comme une occasion inédite pour se confronter à la normativité technique en tant que régime d'activité échappant toujours en partie, dans son fonctionnement concret, à la norme de l'automatisme et recelant une marge d'indétermination (Simondon, 2012) à partir de laquelle il est possible d'envisager une inventivité comportementale à même l'activité sociale.

Il nous faut insister sur deux aspects essentiels pour comprendre en quoi la perspective développée ici se distingue, sans pour autant la réfuter, d'une approche discursive à partir de laquelle la normativité (entendue comme activité productrice de normes) est comprise comme la mise en forme d'une activité comportementale qui serait pure ou amorphe, la formalisation d'une spontanéité comportementale qui existe uniquement comme « dehors » du discours<sup>2</sup>. La première distinction, d'ordre général, est que l'approche com-

---

2. Nous reconduisons ainsi la critique simondonienne du schème hylémorphique selon lequel la matière ne fait que prendre passivement forme. La conception dominante du comportement

portementale ne cherche pas à discréditer l'importance de la construction discursive des normes sociales ; elle vient essentiellement enrichir le critère à partir duquel on peut rendre compte de l'institution, la propagation et la contestation normative. Un énoncé, même compris comme participant d'un système discursif plus général, ne permet pas de rendre compte du « suivi » de la norme autrement que comme une normalisation ; l'énoncé ne contient pas, en plus de lui-même, les conditions extra-discursives pour expliquer comment le suivi de la norme qu'il énonce implique toujours une activité, bien souvent inventive, sur le plan comportemental<sup>3</sup>. La deuxième distinction, qui renvoie plus précisément à notre question, est que le discours a longtemps été pris comme formalisation suffisante pour exprimer les opérations d'une machine, et ceci s'expliquait logiquement par le fait que le discours (en tant que programme) devait précéder l'opération machinique ; le comportement de la machine (l'algorithme) était seulement l'expression d'un ordre discursif préétabli, déjà écrit, déjà énoncé, et tout reste ou excès à ce comportement n'était que le signe d'un dysfonctionnement ou d'une déraison.

Notre proposition cherche à aborder la normativité de l'apprentissage algorithmique à partir du problème du comportement, à considérer le comportement *et* comme dynamique d'invention *et* comme exécution automatique. En disant cela, il faut tenir compte du fait que le comportement *est* une activité à l'intérieur d'une autre activité que l'on appelle ici « social » pour rendre compte de l'effet « encadrant »<sup>4</sup> d'une activité sur une autre. Cet encadrement n'a toute-

---

héritée du behaviorisme continue à le traiter comme une matière, plastique certes, mais qui ne fait que se conformer aux prescriptions discursives et aux contraintes de son environnement, sans voir que cette conformité et cette réactivité supposent elles-mêmes une *activité* et non pas seulement une passivité.

3. Nous nous basons sur l'articulation d'auteurs aussi différents que Georges Canguilhem, Pierre Bourdieu ou Harry Collins pour soutenir une telle proposition, mais Andy Clark en donne une expression très frappante lorsqu'il écrit : « En tant que machines à prédictions sans cesse actives, ce type de cerveau ne s'occupe pas du tout de "traiter des entrées". Cette stratégie neuronale proactive nous prépare à l'action et permet à des agents mobiles et incorporés d'intervenir dans le monde, et de susciter les flux sensoriels qu'ils réclament. [...] Nous ne sommes pas des flans cognitifs qui attendent oisivement la prochaine "entrée" mais davantage des predictavores [*sic*] proactifs – la machine à deviner de la nature qui essaie sans cesse de garder une longueur d'avance en surfant les vagues de stimulations sensorielles entrantes » (Clark, 2016, p. 52, notre traduction).

4. En effet, une autre formulation possible pour rendre compte de ce couplage comportement-activité pourrait être empruntée à Raymond Ruyer qui parle d'activité encadrée et d'activité encadrante (Ruyer, 1954).

fois rien d'unidirectionnel : un comportement peut devenir le foyer d'activité dès lors que son exécution automatique pose problème et réclame une nouvelle organisation<sup>5</sup>. Prendre le couplage *comportement-activité* comme unité d'analyse revient à opérer une suspension du partage entre l'organique et le machinique, non pas pour rendre toute possibilité de partage caduque et pour tomber dans une indifférenciation des modes d'existence, mais précisément afin d'envisager le rapport humain-machine au-delà d'un fonctionnalisme bien établi qui consiste à indexer la machine à l'automatisme et le vivant à l'invention. En effet, la relation entre automatisme et invention est bien plus complexe que le binarisme souvent défendu par la « tortue romaine » de la philosophie qui défend, sur cette base, une ligne de partage entre l'organique et le machinique (Malabou, 2017)<sup>6</sup>. En nous appuyant sur la théorie du

---

5. Harry Collins (1990, p. 22) signale cette réversibilité entre comportement et activité en signalant qu'un même comportement peut signifier différentes activités sociales et qu'une même activité sociale peut se comporter de différentes manières. Pour illustrer cette réversibilité de l'encadrement, prenons l'ouverture d'une porte. Un comportement dont l'exécution pose rarement problème, comme opérer une pression sur une poignée de porte, est généralement conçu comme mécanique dans la mesure où il n'a pas à être *conçu*. Il n'y a pas à se représenter le mécanisme de la porte pour pouvoir l'ouvrir ; l'image de l'ouverture est déjà entièrement intégrée dans un comportement opératoire appris qui permet d'anticiper les effets que nos mouvements auront sur la porte. Mais lorsque cette anticipation est contrariée, c'est l'image de la porte qui se trouble. Notre propre mécanisme et celui de la porte s'enrayent dans un même mouvement et, brusquement, l'ouverture de la porte devient l'activité prégnante. Un comportement « automatique » devient l'enjeu d'une activité qui doit résoudre un problème. En d'autres termes, le mécanisme d'un comportement recèle une marge d'indétermination qui lui permet de se moduler activement en faisant varier l'ordre, la vitesse ou l'intensité des gestes (on dit souvent qu'il faut « jouer avec le mécanisme » de la porte) et de se superposer momentanément à la finalité de l'activité première (rentrer chez soi chargé de courses, quitter une salle de réunion discrètement, etc.) comme problème à surmonter afin de pouvoir reprendre celle-ci. Dans ce cas, l'activité « quitter une salle » n'est pas abandonnée : une fois la porte ouverte, on ne reste pas oisivement dans l'embrasure. Au contraire, le comportement problématique s'intègre à l'activité qui devient alors celle de « pouvoir ouvrir la porte afin de quitter la salle ».

6. Il nous semble par ailleurs que la théorie bourdieusienne de l'*habitus* constitue déjà une des plus importantes tentatives de dépasser cette vieille opposition entre invention et automatisme, et que son renouvellement en dialogue avec les sciences cognitives, l'intelligence artificielle et la philosophie des techniques est un chantier prometteur que nous souhaiterions entreprendre : « Lors même qu'elles apparaissent comme déterminées par le futur, c'est-à-dire par les fins explicites et explicitement posées d'un projet ou d'un plan, les pratiques que produit l'*habitus* en tant que principe générateur de stratégies permettant de faire face à des situations imprévues et sans cesse renouvelées sont déterminées par l'anticipation implicite de leurs conséquences, c'est-à-dire par les conditions passées de la production de leur principe de production [...] » (Bourdieu, 2000).



cycle de l'image de Gilbert Simondon, nous allons ici tenter de montrer que le partage entre machine (automatisme) et organisme (invention) est insuffisant pour rendre compte de la normativité des algorithmes apprenants. Mais d'abord, un détour par la cybernétique s'impose. En effet, le moment où il est devenu possible de concevoir le fonctionnement algorithmique en termes comportementaux, et où en même temps la question du comportement engage directement celle d'une généralisation de la prédiction comme mode de gouvernement, est par excellence le moment cybernétique.

## FEED-BACK, COMPORTEMENT ET PRÉDICTION

Longtemps, l'utilité d'une machine a été évaluée en fonction de sa capacité à obéir à des commandes, à s'exécuter selon l'ordre programmé, écrit à l'avance. La valeur d'une machine tenait à la fois à sa confirmation du discours déterministe de la mécanique classique et à sa manière de représenter docilement le corps productif du travailleur. On n'attendait de la part de la machine aucune surprise ou bifurcation dans son fonctionnement ; celle-ci n'aurait été, à la limite, qu'une valeur négative indiquant un *dys*fonctionnement à corriger, mais pas une invention dont il faudrait prendre acte. Cette conception déterministe de la machine est celle qui irrigue la Révolution industrielle ; le tournant cybernétique entamé au début du XX<sup>e</sup> siècle bouleverse radicalement les attentes et les valeurs accordées à la machine. D'un algorithme comme comportement à exécuter par la machine, on passe alors à l'algorithme comme mécanisme de contrôle et de régulation d'un système imprévisible. L'algorithme ne sert plus uniquement à communiquer à la machine l'ordre écrit d'avance qu'elle doit exécuter, mais il sert, en plus, à *simuler* les phénomènes qui se soustraient à une analyse exhaustive (par exemple l'activité cérébrale). Le problème de la prédiction émerge justement lorsque la logique formelle d'un programme ne suffit plus à rendre compte de la dynamique de fonctionnement d'une machine. C'est alors que cette dernière rompt avec l'idée d'un mécanisme rationnel et transparent, parfaitement adéquat au discours qui l'ordonne, pour prendre les allures d'une boîte noire dont on ne peut tout connaître et qui se soustrait en partie à l'ordre du discours.

L'exemple cybernétique classique d'un mécanisme prédictif vient de la balistique, notamment le système de tir antiaérien qui doit prédire la trajectoire de sa cible afin que son tir la devance. Pour ce faire, il faut extrapoler statistiquement la trajectoire *probable* de l'avion. Tirer sur un avion est d'abord et avant tout un problème de prédiction. L'extrapolation statistique de la trajectoire de

l'avion doit produire un résultat dans une fenêtre temporelle très restreinte, sans quoi la prédiction émise sur la base des données en entrée ne sera plus valable. La prédiction est la plus précise au temps  $t$ , chaque instant qui passe ne faisant qu'éloigner la prédiction de sa cible virtuelle. Ou plutôt faudrait-il dire que *le nombre de prédictions possibles s'accroît à mesure que le temps s'écoule, et que par le même coup, la prédiction émise recouvre de moins en moins l'espace des positions possibles*. De ce point de vue, deux choses deviennent essentielles : la diminution du laps de temps qui s'écoule entre l'entrée et la sortie du processus, et une rétroaction permanente entre la prédiction et les entrées qui la nourrissent. La cybernétique nous présente ainsi un modèle où la prédiction et l'action ne sont plus extérieures l'une à l'autre, mais sont corrélées au sein d'une même activité téléologique, se nourrissant mutuellement à travers des boucles de rétroaction. D'un point de vue ontologique, il faut bien insister sur le fait que cette conception de la prédiction et de l'action ne se limite pas forcément à un travail expert de modélisation, mais est une caractéristique transversale de tous types d'activités (neuro-physiologique, consciente, sociale, etc.). Dans un système où la stabilité doit sans cesse être regagnée, le contrôle devient synonyme de prédiction et d'approximation des comportements du système. Sans doute n'avons-nous jamais pleinement saisi la fonction presque existentielle que jouent l'erreur et l'échec dans la théorie cybernétique du contrôle et de la communication. Dans une série temporelle statistique, l'erreur est le moteur de la rétroaction ; avant la répétition du même il y a sa mise en échec imparable. Et sans doute devons-nous comprendre la généralisation de la prédiction à travers le *machine learning* comme étant intrinsèquement liée à ce reste inlassable d'erreur dans tout système dynamique.

Norbert Wiener insiste sur cette prévalence de l'erreur comme signe localisé de l'entropie dans un système de communication. Il accorde une primauté à l'écart, à la différence entre l'entrée et la sortie. Après l'exemple cybernétique du gouvernail de navire (dont l'étymologie gréco-latine infuse le terme cybernétique), il prend celui, d'apparence plus banale, de la prise en main d'un crayon. À chaque instant de cette opération, la quantité de prise en main du crayon qu'il reste encore à assurer décroît sans que l'on puisse pour autant déterminer avec certitude le moment à partir duquel le crayon est effectivement pris en main : « La performance d'une telle action suppose un rapport permanent, conscient ou inconscient, vers le système nerveux qui indique la marge avec laquelle nous avons échoué à prendre en main le crayon à chaque instant » (Wiener, 1961, notre traduction). C'est exactement ce que relève aujourd'hui le philosophe de la cognition Andy Clark lorsqu'il précise que,

si la prédiction renvoie, dans le langage courant, à un acte conscient ou subjectif d'anticipation dans une situation d'incertitude, le vrai défi du point de vue comportemental est de comprendre comment l'action est en permanence orientée par des processus prédictifs « inconscients » et « automatiques ». « Ce type de prévoyance, déployée automatiquement, profondément probabiliste et non conscient qui se produit au niveau des processus complexes de traitement neuronal, sous-tend et unifie la perception et l'action » (Clark, 2016, p. 2, notre traduction). On voit bien ici que la prédiction ne se réduit pas seulement à une production discursive dont l'algorithme serait la forme, mais qu'elle émane de l'activité comportementale.

Dans le célèbre article de 1943 « Behavior, Purpose, and Teleology », coécrit avec le physiologue Arturo Rosenblueth et l'informaticien Julian Bigelow, Wiener définit le comportement (*behavior*) comme étant tout changement d'une entité par rapport à son milieu : « Il est possible que ce changement soit, en grande partie, la sortie [*output*] d'un objet, l'entrée [*input*] étant minimale, distante ou sans pertinence ; ou alors le changement peut être immédiatement indexé à une certaine entrée. Il en suit que n'importe quelle modification d'un objet qui est détectable extérieurement est susceptible de s'appeler comportement » (Rosenblueth *et al.*, 1943, p. 18, notre traduction). À en rester à une définition aussi englobante du comportement, les auteurs reconnaissent que celui-ci risque bien de se vider de lui-même, raison pour laquelle ils entreprennent une typologie en cascade des différents types de comportements à partir d'une distinction assez classique entre un *comportement actif* – l'objet est la source de l'énergie actuelle (*output*) dans une réaction donnée – et un *comportement passif* – l'objet ne fait que traduire une énergie incidente (*input*) en énergie actuelle. Le comportement actif peut ensuite être *purposeful* ou *purposeless*, avec ou sans objectif, détermination. Dans le premier cas, l'orientation corrélée à celle d'un autre objet conditionne le comportement, c'est le cas par exemple d'une torpille dotée d'un servomécanisme de ciblage (Rosenblueth *et al.*, 1943, p. 19). Dans le second cas, le comportement n'est orienté vers aucun état final ; c'est le cas de l'horloge dont la fonction n'explique en rien son comportement. La notion de détermination étant acquise, la distinction suivante dans leur typologie s'opère au niveau d'un « *purposeful active behavior* » qui peut lui-même être subdivisé en deux classes : *rétroactif (feed-back)* ou *téléologique et non rétroactif (non-feed-back)* ou *non téléologique*. Cette distinction va permettre d'établir le type de relation qu'entretient le comportement avec son objectif – qu'il faut de nouveau comprendre comme une orientation qui s'exprime à même le comportement. Un processus de rétroaction peut être positif – la sortie est reprise à l'entrée dans un

mouvement d'amplification du signal – ou négatif. Dans cette dernière instance, « le terme rétroaction [*feed-back*] est employé de manière plus restrictive pour signifier que le comportement d'un objet est contrôlé par la marge d'erreur dans laquelle se trouve l'objet à un *temps* donné en fonction d'un *objectif* relativement spécifique » (Rosenblueth *et al.*, 1943, p. 18, notre traduction, nous soulignons).

Cette qualification négative de la rétroaction est la base à partir de laquelle ils continueront à décliner leur typologie du comportement sous ses aspects *prédictif* ou *non prédictif*. Certaines pathologies affectant les fonctions proprioceptives et motrices impliquent des phénomènes de rétroactions positives dans lesquels il n'y a pas de correction temporelle du signal nerveux au fur et mesure où le comportement rapproche l'objet de son objectif. C'est le cas de patients porteurs de syndromes neuro-dégénératifs dont les tremblements vont en s'amplifiant lorsqu'ils approchent le verre d'eau de leur bouche. Dans ces cas, le système nerveux succombe à des oscillations amplifiantes qui finissent par détraquer le comportement et par entraver l'accomplissement de l'objectif. De manière imagée, on pourrait dire que le *feed-back* positif consiste à « en faire trop »<sup>7</sup>. En revanche, les mécanismes de rétroaction négative assurent une correction permanente du comportement en fonction de sa « réussite » partielle. En ce sens, un comportement téléologique est un comportement contrôlé qui obéit à une fin ; les montages moteurs sont des analogues de servomécanismes. On comprend mieux dès lors l'exemple de la saisie du crayon : celle-ci implique un retour permanent de son relatif échec à chaque instant du comportement en cours. Il faut immédiatement insister sur le fait que la téléologie n'est pas ici à comprendre dans les termes d'un schème causaliste comme un acte obéissant à une cause finale. En indexant *la téléologie à un objectif qui se présente comme le comportement prédictif lui-même*, les auteurs entendent rendre compte de comportements qui se corrigent à partir de leur marge d'erreur, c'est-à-dire la différence entre l'état de l'objet en train de se comporter à un moment donné et l'état final interprété comme l'objectif de ce comportement (Rosenblueth *et al.*, 1943, p. 24). En prenant le comportement comme unité d'observation et d'analyse de systèmes complexes, la cybernétique permet d'envisager la machine comme une *activité* et non plus uniquement comme un objet fonctionnel délimité. Malgré cette ouverture, la conception de cette activité reste largement empreinte d'une

---

7. À ce titre, l'origine balistique de l'expression analogue souvent utilisée en anglais « over-shooting the target » est significative.

certaine conception régulatrice, et peine à concevoir en quoi le déploiement de comportements machiniques pourrait être l'occasion d'une réinvention du sens même de l'activité.

Avant d'en revenir à notre objet, à savoir le caractère prédictif du *machine learning*, et dès lors qu'on prend le comportement et l'activité comme unité d'analyse, nous aimerions proposer une perspective alternative à celle de la cybernétique pour rendre compte du caractère extensif et non réductionniste de la prédiction. Cette alternative, nous la trouvons dans la philosophie de l'individuation de Simondon, et plus particulièrement dans sa théorie du cycle de l'image qui permet d'envisager le couplage autrement que comme un rapport d'exécution ou de mise en forme de l'un par l'autre.

## CYCLE DE L'IMAGE ET MARGE D'INDÉTERMINATION

À bien des égards, la philosophie de Simondon résonne avec la métaphysique cybernétique. Elle s'en distingue cependant radicalement par sa conception de l'information. Là où la cybernétique voit l'information comme une fonction de régulation et de stabilité, Simondon la comprend comme une fonction disruptive qui induit des transformations au sein d'une structure qui recèle une marge d'indétermination. L'information est une opération qui permet à une structure d'en devenir une autre ; elle est la dynamique du devenir. De ce point de vue, les individus sont toujours les termes limites et passagers d'un processus, d'une genèse, et non pas l'unité à laquelle le divers, le flux, le devenir s'indexerait. Pour approfondir, prenons une pente particulière de la philosophie simondonienne, à savoir sa théorie du « cycle de l'image » telle qu'elle est développée dans son cours *Imagination et invention*. En réitérant le geste novateur de sa philosophie de l'individuation, Simondon montre, à travers une problématisation attentive des travaux en psychologie et éthologie de son époque, que l'imagination n'est pas simplement une activité de production d'images, elle indique aussi la réception et l'accueil d'images – lesquelles ont par ailleurs une certaine autonomie par rapport à toute position ou intention perceptive. En fait, l'image mentale qui a une signification pour une conscience et l'image qui est l'objet d'une perception sont deux cas limites d'une *activité imaginaire* qui passe par des phasages et des déphasages successifs au long d'un processus de développement (Simondon, 2008, p. 3). Avant même de rencontrer un objet, un corps tente des mouvements, anticipe son milieu, suscite par son activité la rencontre d'une extériorité et l'actualisation de schèmes moteurs. L'expérience de cette extériorité suscite un *problème* au

niveau des anticipations perceptivo-motrices, problème qui exige une réorganisation interne du schème, l'*invention* d'une nouvelle formalisation ou systématisation pouvant intégrer son expérience ; formalisation qui prépare de nouvelles anticipations du milieu, de nouveaux montages perceptivo-moteurs. Voilà les trois phases du cycle de l'image : anticipation, expérience, systématisation. Ces niveaux ou phases correspondent respectivement à la charge motrice, cognitive et affectivo-émotive des images : *i)* le premier niveau renvoie à l'engagement corporel comme activité – le fait que l'organisme, avant même d'opérer une différenciation objectale, est pris tout entier dans une relation avec son milieu, grâce auquel il actualise des schèmes moteurs ; *ii)* le deuxième niveau permet à l'organisme de se libérer de l'engagement total dans son milieu d'activité, dans la mesure où il « développe un analogue mental de ce rapport primaire » (Simondon, 2008, p. 22) et peut donc se représenter un objet stabilisé et reconnaissable d'une situation à un autre ; *iii)* le troisième niveau est la phase de systématisation par laquelle un sujet se conçoit réflexivement comme ce par rapport à quoi le monde s'organise de manière appréhendable par des logiques d'ordre supérieur. L'invention quant à elle agit comme la relance du cycle, elle en assure le dynamisme, un dynamisme qui ne se limite pas à une fonction de rétroaction négative, mais qui correspond à l'ouverture productrice de nouvelles normes.

La théorie de l'imagination productrice que Simondon esquisse, entendue comme activité de production d'images, cherche à maintenir celle-ci à la fois comme activité de représentation *et* d'invention. L'activité n'est pas simplement le cadre dans lequel un sujet et un objet préconstitués se mettent en relation ; c'est à l'inverse l'activité qui est le foyer de la relation à partir de laquelle il est possible de parler d'une polarisation continue entre un sujet et un objet. Or, au sein d'une activité donnée, c'est une pluralité d'ordres de grandeur qui doivent communiquer de sorte que la divergence de leurs normes puisse se constituer en un régime d'activité commun, quoique précaire et passager. Comme Simondon l'écrit dans *Imagination et invention* :

« [C]es schèmes d'action existent donc dans l'être vivant comme anticipation des conduites possibles, comme programmes partiels des comportements, et ils peuvent, virtuellement utiliser, fournir un contenu aux anticipations, sous forme de préparation de situations de rencontre de l'objet et d'anticipation des réponses ; l'organisme peut, plus ou moins complètement, jouer à vide ses conduites avant de les appliquer à un objet réel ; se lever, attaquer, se cacher, fuir, faire face, ce sont des séquences dont l'être vivant possède le programme en lui-même, comme il possède son propre corps ; il y a là une base pour

l'activité locale d'anticipation, qui peut aller jusqu'à un jeu moteur gratuit, les situations étant improvisées, mais le fondement de l'activité étant fourni par ces schèmes que l'être possède et qu'il peut à tout moment susciter, comme s'il était dans une situation » (Simondon, 2008, p. 32)<sup>8</sup>.

Concevoir un schème d'action comme un « programme partiel des comportements » implique qu'une activité ne soit pas simplement un déroulement ou une séquence de comportements en attente d'activation – comme s'il pouvait exister des montages comportementaux tout faits, mobilisés pour telle ou telle action –, mais que le montage demeure partiellement incomplet, ou plutôt qu'il se complète en acte. L'action donne sa finalité et son sens au montage comportemental après coup. Il n'y a pas d'abord une représentation de la finalité de l'action et ensuite le comportement qui vient l'instancier, mais plutôt une dépense comportementale permanente qui trouve une certaine unité dans le sens de l'action. Ce sens n'est jamais tributaire de la seule échelle individuelle, mais toujours de la dimension collective de l'activité. En ce sens, on peut dire qu'il existe un *répertoire de comportements potentiels* à toute activité, répertoire qui est tout autant affaire d'individu que de milieu<sup>9</sup>.

Notre hypothèse est que les comportements participant au *machine learning* se déploient, eux aussi, à travers des phases d'anticipation, d'expérience et de formalisation. Avec le *machine learning*, la statistique, en plus d'être un outil ou une technique de savoir, une certaine mise en forme discursive ou un régime de vérité, devient de surcroît une activité sociale se déployant à travers des comportements automatiques *et* inventifs, sans que l'on puisse statuer *d'avance* si c'est une machine qui exécute des comportements automatiques ou si c'est un organisme qui invente des comportements. L'important est de souligner que le partage entre l'automatisme et l'invention n'est pas exclusif, mais doit se comprendre comme une dynamique de phases dans un cycle : il y a donc une certaine réversibilité entre les deux termes selon le sens de

---

8. Ce passage frappe par sa résonance avec la définition bourdieusienne de l'habitus et en particulier avec certains passages où Bourdieu explicite plus clairement ce qu'il entend par habitus : « Lors même qu'elles apparaissent comme déterminées par le futur, c'est-à-dire par les fins explicites et explicitement posées d'un projet ou d'un plan, les pratiques que produit l'habitus en tant que principe générateur de stratégies permettant de faire face à des situations imprévues et sans cesse renouvelées sont déterminées par l'anticipation implicite de leurs conséquences, c'est-à-dire par les conditions passées de la production de leur principe de production. [...] » (Bourdieu, 2000).

9. Pour une présentation approfondie de cette notion de « répertoire de comportements potentiels », voir Reigeluth (2018).



l'activité sociale. C'est ce qui fait dire à Harry Collins qu'un même comportement peut signifier plusieurs activités et qu'une même activité peut se déployer à travers divers comportements (Collins, 1990, p. 33). D'ailleurs, selon Collins, l'intelligence artificielle n'est pas une prothèse organique, mais bien une prothèse *sociale* (Collins, 1990, p. 14). Ceci nous permet alors d'ouvrir la perspective de Desrosières à de nouveaux enjeux. Dans *La politique des grands nombres*, par exemple, celui-ci montre que les « outils cognitifs de généralisation » (Desrosières, 2010, p. 395) que sont les techniques statistiques reposent sur et construisent des espaces d'équivalences discursives à l'intérieur desquels des objets disparates sont rendus commensurables. Or il semblerait que cette équivalence du discours statistique ne soit plus uniquement redevable à une norme discursive, et que la généralisation opérée est autant le produit des comportements machiniques qu'elle l'est des comportements humains. *Longtemps liée à l'activité et à la participation du statisticien ou du scientifique, l'individuation des schèmes des techniques statistiques doit, avec le machine learning, compter parmi ses participants l'activité de la machine elle-même.* Ceci ne veut pas dire que la machine remplace ou automatise forcément le savoir-faire de l'agent humain, mais simplement que, pour rendre compte du fonctionnement concret des algorithmes apprenants, il faut y inclure leurs comportements en ce qu'ils participent à, et sont encadrés par, une activité. Clarifions à présent ce dont il s'agit.

### LE MACHINE-LEARNING : REPRÉSENTER, ÉVALUER, OPTIMISER

L'enjeu central de l'apprentissage algorithmique est de chercher la fonction qui pourra le mieux généraliser les occurrences des données sous forme d'une prédiction. Comme le souligne Mackenzie, il existe différents *styles* de prédiction :

« En décrivant comment les techniques de *machine learning* trouvent des fonctions, il ne s'agit pas de dire que nous devrions avoir une saisie détaillée de leur fonctionnement. Je suggère plutôt que nous devons nous confronter aux différences entre les processus de trouvaille de fonction [*function-finding*] associés à certains contextes de *machine learning* et désirs prédictifs. Nous devrions distinguer entre des styles prédictifs. Même du point de vue d'une économie politique relativement simple, la valeur des prédictions diffère en fonction du travail qui les produit, tout comme différents styles prédictifs (probabiliste, théorico-informationnel, décisioniste) impliquent différents types de valeur » (Mackenzie, 2015, p. 436, notre traduction).



On entrevoit ici que le *machine learning* – en tout cas dans ses formes non supervisées les plus poussées communément appelées *deep learning* – est non seulement mobilisé et déployé à l'intérieur de certaines activités sociales, mais qu'il est de surcroît en lui-même une activité qui fait tenir ensemble des comportements disparates d'une certaine *manière*. L'idée n'est donc pas de revenir à une position relativement classique qui considère la normativité technique comme l'expression de valeurs sociales, mais au contraire de faire une place pour la marge d'indétermination des comportements algorithmiques, de sorte qu'on puisse se poser de nouvelles exigences épistémologiques et libérer de nouvelles potentialités d'action. Essayons de dégager les phases structurantes de cette activité.

Le premier défi qui se présente à la construction des modèles prédictifs du *machine learning* concerne le choix de la représentation grâce à laquelle le modèle pourra accéder aux données déjà organisées. L'enjeu principal de la représentation est de se situer dans la famille algorithmique (connexionniste, symboliste, évolutionniste, bayésienne, analogique) qui va orienter les modèles, techniques et méthodes utilisés. C'est sans doute ici que la discursivité garde encore sa plus grande prégnance en tant que mise en forme conventionnelle des données dans la mesure où le choix de la représentation revient en quelque sorte à adopter une *vision du monde* qui décide non seulement des critères de représentation à partir desquels un phénomène se construit, mais qui prépare également le type d'évaluation et d'optimisation de l'algorithme qui sera valorisé. Il s'agit de déterminer le *langage* grâce auquel le modèle va accéder au monde de données structurées. Il faut se garder du fantasme selon lequel un algorithme établirait de l'ordre ou inventerait une formalisation à partir de l'existence « brute » et désordonnée des données (Gitelmann, 2013). Ces dernières sont toujours déjà structurées à la fois par le dispositif technique qui les produit et les enregistre, et par les pratiques de traitement et de « nettoyage » informatiques par lesquelles elles passent nécessairement pour convenir à la forme de représentation du modèle. Elles sont toujours déjà embarquées dans une activité sociale. Loin d'être une matière amorphe, les données ont déjà leurs formes implicites qui les prédisposent à certaines informations. « Bien souvent, les données brutes ne se présentent pas dans une forme qui facilite l'apprentissage, mais il est possible de construire des variables [*features*] qui le sont. » (Domingos, 2015a, notre traduction). Les théoriciens du *machine learning* insistent par ailleurs sur l'étonnante disproportion de temps nécessaire à la récolte, à l'intégration et au nettoyage des données, par rapport à la phase tout aussi pénible d'essais et de corrections dans l'ingénierie des variables (*features*)

(Flach, 2012, p. 13). Mackenzie précise qu'à bien des égards, les *features* du *machine learning* jouent essentiellement le même rôle que les variables dans la statistique classique, si ce n'est que les *features* peuvent provenir de n'importe quelle source de données (texte, vidéo, images, sensors, etc.), là où les variables correspondent typiquement à un travail de tabulation de sondages, d'enquêtes ou d'échantillons aléatoires (Mackenzie, 2015, p. 433). En ce sens, les *features* n'expriment pas nécessairement des relations de dépendance ou d'influence logiques comme les variables, mais simplement ce qui est susceptible d'émerger comme forme prégnante au sein d'un jeu de données à un temps  $t$ . Il est important de souligner que cette forme est une certaine expression d'une représentation et qu'elle n'a rien de spontané. Elle est, dans une certaine mesure, l'anticipation générée et partiellement déterminée par une formalisation antérieure.

La représentation recèle alors une dimension sémantique et syntaxique ; le choix du langage informatique va déterminer la manière dont l'algorithme communique avec l'ordinateur alors que le choix des *features* va déterminer le degré de dimensionnalité des données à traiter. Selon Flach, le modèle, en *machine learning*, ne correspond à rien d'autre qu'au type de connexion qui va lier les *features* ensemble. La dimensionnalité est un problème incontournable dans la construction d'un modèle, à tel point qu'il a été baptisé par les théoriciens et praticiens du *machine learning* « la malédiction de la dimensionnalité » (*curse of dimensionality*) (Marsland, 2009, pp. 17-18 ; Domingos, 2015b ; Flach, 2012 ; Witten et Frank, 2011). Cette formule désigne le fait qu'il devient exponentiellement compliqué de généraliser correctement à partir d'une quantité de variables grandissante, étant donné qu'un *jeu de données d'entraînement* (la part de données qui est utilisée à titre d'exemple pour générer/apprendre des classes de données) recouvre une fraction décroissante de l'*espace d'entrées* (l'ensemble des données que le modèle est censé représenter). Pour conserver sa disposition à généraliser et à classer les occurrences en fonction des variables, il faudra que l'algorithme se nourrisse toujours de plus de données. Le temps, comme pour n'importe quel problème effectivement calculable, demeure une ressource élémentaire dans l'apprentissage algorithmique : « D'immenses montagnes de données sont disponibles, mais il n'y a pas assez de temps pour les traiter. Elles restent alors inutilisées. Ceci mène à un paradoxe : bien qu'en principe la disponibilité de plus de données signifie que des classificateurs plus complexes peuvent être appris, en pratique on finit par utiliser des classificateurs plus simples étant donné qu'il faut moins de temps pour les apprendre » (Domingos, 2015b, notre traduction). La « malédiction » du modèle algorithmique est donc de grandir ou périr, et

le temps qui lui est imparti pour le faire sera toujours limité. Ceci introduira, comme on le verra ci-après, la question de l'optimisation.

Si les algorithmes apprenants sont si souvent analysés en termes comportementaux, c'est que leur fonctionnement effectif ne peut s'expliquer uniquement par leurs paramètres ou règles de programmation initiale. La prédiction à partir d'une généralisation est toujours, *dans une certaine mesure*, en proie à l'échec. La part de cet échec est précisément ce qui va faire l'objet d'une évaluation ; c'est là le sens même d'une probabilité en statistique. Il nous faut élucider deux enjeux statistiques pour mieux comprendre ce sur quoi porte exactement l'évaluation : le biais et la variance. « Le biais est la tendance qu'a l'apprenant d'apprendre la même mauvaise chose. La variance est la tendance d'apprendre des choses aléatoires, sans tenir compte du signal réel » (Domingos, 2015b, notre traduction). L'enjeu du *machine learning* est d'apprendre la fonction lui permettant d'établir des relations entre les données de sorte à pouvoir prédire les occurrences futures sur la base de ces relations. Il va de soi que la marge d'erreur s'évalue par rapport à la tâche qu'il est demandé à l'algorithme d'apprendre à réaliser (classification, régression, recommandation, etc.), mais l'alternative *biais-variance* reste une préoccupation transversale à l'apprentissage supervisé qui va s'exprimer au travers des différentes techniques que la tâche en question exige. Un biais appauvrit l'exactitude<sup>10</sup> du modèle et fait qu'il ne colle pas aux données « réelles », ce qu'on appelle un « sous-ajustement ». La variance, en revanche, octroie trop d'importance à la variabilité des données, ce qui tend à fragmenter le pouvoir généralisateur du modèle, phénomène appelé « sur-ajustement ». Généralement, un modèle plus complexe, c'est-à-dire avec un degré de dimensionnalité plus élevé, aura tendance à réduire le biais et à augmenter la variance. Toutefois, si les effets de ce compromis ne sont pas minimisés, on aura affaire à un modèle certes puissant en matière de complexité, mais complètement impuissant en ce qui concerne la généralisation. « Ainsi, contrairement à l'intuition, un apprenant plus puissant n'est pas forcément meilleur qu'un apprenant moins puissant » (Domingos, 2015b, notre traduction). Un équilibre entre le sous-ajustement et le sur-ajustement doit être trouvé (Flach, 2012, p. 93).

Il existe alors au moins deux types d'erreurs : l'une – le biais – dérivée des hypothèses erronées à la base du modèle qui force une généralisation trop

---

10. Nous traduisons « accuracy » par « exactitude » qui correspond à la somme des vrais négatifs et vrais positifs effectués par la classification, divisée par le nombre total d'exemples. Voir Marsland (2009).

grossière ; l'autre – la variance – liée à la variabilité statistique de l'échantillon d'entraînement qui entraîne un modèle à tendance particulariste. En d'autres mots et qu'importe son origine, l'erreur est à la fois contrôlable et irrépessible dans toute fonction d'apprentissage. La marge d'erreur dans les prédictions du modèle est décomposée en une mesure du biais et une mesure de la variance. Cette décomposition ne suffit pas pour autant à rendre l'erreur analytiquement localisable, elle devra toujours prendre en compte un troisième terme : le « bruit », ou l'erreur irréductible, qui correspond à la variance du jeu de données tests, c'est-à-dire les données sur lesquelles le modèle devra produire des prédictions plus ou moins exactes (Marsland, 2009, p. 36). L'évaluation de l'erreur est avant tout une décomposition permettant, dans une certaine mesure, de localiser la source et de caractériser la forme de l'erreur. Comme la main de Wiener qui ne peut apprendre à saisir le crayon que s'il ne l'a jamais entièrement saisi, la puissance prédictive d'un modèle passe par un échec relatif et contrôlé à s'adapter *complètement* aux données sur lesquelles il s'entraîne. La dimension d'évaluation de l'apprentissage algorithmique consiste en quelque sorte à déterminer quel serait le comportement le plus adapté à la tâche donnée, tout en sachant que la généralité de la tâche ne peut prévoir toutes les particularités des problèmes à résoudre (Flach, 2012, p. 13). Il est important d'insister ici sur le fait que la mesure de l'« erreur » ou de l'« échec » d'une prédiction d'un modèle n'est pas un jugement sur sa « vérité ». On cherche avant tout des modèles qui *font* quelque chose, qui accomplissent une tâche, qui sont utiles. C'est une leçon cybernétique qui se rejoue : le comportement est orienté vers un objectif sans forcément être objet de conscience. Raison pour laquelle aussi certains classificateurs sont préférés à d'autres précisément parce que leurs hypothèses sont tellement « fausses » qu'elles permettent de réduire la marge d'erreur due à la variance<sup>11</sup>.

Si la généralisation et la prédiction sont les fonctions fondamentales du *machine learning*, alors un paradoxe surgit immédiatement : le comportement à optimiser ne peut être déterminé qu'après coup. Contrairement aux problèmes d'optimisation classiques liés à la planification du chemin le plus court sur la base d'une computation des combinatoires possibles, ici la fonction à optimiser n'est pas donnée à l'avance ; tout l'intérêt de la prédiction est

---

11. C'est le cas du classificateur bayésien naïf qui opère selon l'hypothèse d'une indépendance factice des probabilités lui permettant d'obtenir pour certaines tâches des résultats bien plus exacts en termes prédictifs que d'autres modèles concurrents.

d'être effective sur des données inédites, ce qui implique que le modèle n'a souvent aucune représentation – ou qu'une représentation très partielle – de ce que serait l'optimum global du réseau. Il faut *trouver* la fonction. Comme le souligne d'ailleurs Domingos, « remarquez que le fait que la généralisation soit l'objectif implique une conséquence intéressante pour le *machine learning*. Contrairement à la majorité des problèmes d'optimisation, nous n'avons pas accès à la fonction que nous voulons optimiser ! » (Domingos, 2015b, notre traduction). L'algorithme d'optimisation consiste à faire converger dans le temps les différentes fonctions distribuées et parallèles de l'apprentissage vers une fonction d'apprentissage générale qui confère au modèle un point d'équilibre dynamique : la structure du modèle est liée à la manière dont elle corrèle ses prédictions à ses observations ; l'idée étant que la fonction d'apprentissage doit de mieux en mieux prédire les observations, mais aussi que sa véritable force réside dans sa capacité à s'adapter à tout changement dans les *patterns* d'observations. En ce sens et pour le dire dans des termes cybernétiques, un algorithme apprend lorsque ses prédictions convergent avec ses nouvelles observations. Le comportement algorithmique ne peut être réduit à un simple enchaînement de causes ou à une combinatoire plus ou moins complexe. Domingos insiste d'ailleurs sur le fait que « le *machine learning* n'est pas le behaviorisme. Les algorithmes apprenants modernes sont capables d'apprendre des riches représentations internes et pas seulement des associations entre stimuli » (Domingos, 2015a, p. 37, notre traduction). La fonction d'apprentissage optimale n'est pas, comme c'est le cas avec les problèmes de planification, en concurrence avec la preuve mathématique formelle selon laquelle aucun algorithme d'apprentissage ne peut surpasser une fonction de sélection stochastique aléatoire qui essaierait toutes les fonctions possibles. L'apprentissage algorithmique n'est pas une question d'atteindre un optimum fixé et formalisé à l'avance, mais d'*induire* un optimum qui ait une valeur pour une activité située. *La formalisation doit émerger de l'activité et non pas l'inverse*. Domingos va aussi dans ce sens lorsqu'il écrit que le *machine learning* épouserait une approche inductive et observationnelle plutôt qu'une méthode déductive expérimentale :

« Très souvent, *apprendre des modèles prédictifs a pour objectif de les utiliser pour guider l'action*. Si nous découvrons que les langes et la bière sont souvent achetés ensemble au supermarché, alors le fait de placer la bière à côté des langes augmentera sans doute les ventes. [...] Le *machine learning* s'applique en général à des données *observationnelles* où les variables prédictives ne sont pas contrôlées par l'apprenant, contrairement aux données *expérimentales* où elles le sont » (Domingos, 2015b, notre traduction).

Ceci ne doit en aucun cas nous amener à minimiser l'importance de la fonction objective dans l'apprentissage supervisé, qui agit bien, elle, comme une norme discursive, comme une métrique relativement stabilisée, à partir de laquelle les comportements vont être corrigés. Le fait que la maximisation de cette fonction objective soit si souvent liée à une « action » dans un contexte marchand, comme l'indique l'exemple de Domingos, ne doit pas être occulté. L'« optimum » du comportement algorithmique est fonction de ce que l'on choisit d'observer. La cybernétique nous a appris qu'il est impossible de véritablement dissocier la prédiction et l'action, elles sont simplement deux phases d'un même processus : l'action est toujours au-devant d'elle-même, elle est déjà la prédiction en acte, ce qui permet de dire réciproquement que la prédiction est interne à l'action. La prédiction est l'action qui se représente à elle-même son orientation. En ce sens, les comportements d'apprentissage des algorithmes ne feraient pas autre chose que modéliser ou imiter le caractère prédictif de nos propres comportements. Il semblerait qu'on doive définitivement faire le deuil d'une complétude informationnelle, d'une transparence de l'action par rapport à elle-même, de l'idée que l'action découlerait d'un plan. Cela ne fait que révéler le fait que de toute part et en tout temps nous agissons selon des prédictions dont nous n'avons même pas conscience, nous découvrons sans cesse après coup le sens de ce que nous faisons. Comme le soulignent Brian Christian et Tom Griffiths dans leur ouvrage *Algorithms to Live By* : « La norme intuitive d'une prise de décision rationnelle est de soigneusement prendre en compte toutes les options et de choisir la meilleure. À première vue, les ordinateurs seraient les parangons de cette approche en moulinant des computations complexes jusqu'à ce qu'ils arrivent à la réponse parfaite. Mais [...] c'est une image galvaudée de ce que font les ordinateurs : c'est un *luxe* que seul un problème facile permet » (Christian et Griffiths, 2016, p. 261, notre traduction). Mais, toujours selon eux, les problèmes pour lesquels le *machine learning* constitue une plus-value sont des problèmes « compliqués » qui requièrent des réponses moins qu'optimales, mais efficaces et rapides, et surtout qui permettent de « faire mieux qu'avant », ne fût-ce que marginalement. Tout comme notre quotidien où nous prenons sans cesse des décisions sur la base d'informations partielles et où nous agissons régulièrement sans avoir la certitude des effets produits par notre action (Christian et Griffiths, 2016, p. 262).

La prédiction automatisée – ce que ferait le *machine learning* à l'image de nos propres comportements – va de pair avec une nouvelle physique sociale qui semble cependant explicitement récuser la valeur et la signification de la détermination sociale des comportements si chère à Comte ou même à Quetelet.

L'idée qu'un comportement pourrait s'*expliquer* par des *causes* sociales est un non-sens pour les tenants de la nouvelle physique sociale, tel Alex Pentland, pionnier des algorithmes et guru du MIT. Outre ses nombreuses publications scientifiques – qui portent notamment sur les méthodes bayésiennes<sup>12</sup> –, il a également publié deux ouvrages *Honest Signals* et *Social Physics* dans lesquels il loue les apports inédits des dispositifs qu'il participe à inventer dans les laboratoires du Human Dynamics Lab du MIT et qui permettent d'analyser les comportements de groupes. L'inversion de la perspective par rapport à l'attitude sociologique qui se développe à partir du germe de la physique sociale du XIX<sup>e</sup> siècle est claire. Ici, il s'agit non pas de rendre compte de la signification sociale d'un comportement individuel sur la base de déterminations structurales, mais plutôt de faire découler la structure du groupe de la dynamique de ses comportements, d'une classe statistique de comportements. Plus de place ici pour les classes sociales ou les explications sociologiques. Toutes nos interactions sont saturées par des signaux infra-conscients qui sous-tendent l'« intelligence collective » du groupe. « Étant donné que les gens ont cette capacité de signalisation et de réaction automatique, les signalisations de chaque personne peuvent se propager à travers les chaînes de personnes qui constituent leur réseau social, affectant à terme le comportement du groupe dans son entièreté » (Pentland, 2010, p. 38, notre traduction). Ces signaux seraient « honnêtes » précisément parce qu'ils n'émaneraient pas d'une interprétation réflexive de nos propres actions, mais directement des comportements qui génèrent l'action. Une chose est certaine, l'idée d'une raison réfléchissante chère aux Lumières est définitivement enterrée (Pentland, 2010, p. 88). À sa place on retrouve une véritable « technologie du comportement » telle que l'avait appelée de ses vœux le behavioriste Skinner dans son *Beyond Freedom and Dignity* (Skinner, 1971, p. 10).

Adossés à la nouvelle « science des données », des prophètes comme Pentland ont commencé à imaginer cette technologie et à lui donner corps. « Le fait de pouvoir continuellement et universellement mesurer le comportement humain

---

12. Dans une logique bayésienne, une prédiction revient essentiellement à émettre des probabilités *a posteriori* qu'un événement se produise sur base des fréquences des événements observés, et s'il se produit il devient la nouvelle probabilité *a priori* de son occurrence. Il n'est pas étonnant que les systèmes de prédiction reposent souvent sur des modèles bayésiens comme c'est également le cas dans la détection automatique de fraude ou de triche : Marks P., « *The AI that learns our habits and knows when people cheat* », BBC Future, <http://www.bbc.com/future/story/20170126-the-ai-that-knows-when-you-cheat-by-learning-your-habits>, consulté le 6 septembre 2017.



nous fournira la capacité d'organiser [*to engineer*] nos vies comme jamais auparavant. L'espoir est qu'avec cette combinaison de capteurs à la sociométrie et de modèles computationnels du comportement, nous pourrions produire une société sensible – une société qui s'accorde davantage à nos idéaux et aspirations » (Pentland, 2010, p. 91, notre traduction). Pour le dire autrement, l'homologie sous-jacente qui permet de penser la validité et l'intérêt heuristique des « signaux honnêtes » est celle posée entre l'algorithme et le comportement : « C'est parce que nous avons des comportements qui fonctionnent comme des algorithmes que des algorithmes peuvent se comporter de façon à les révéler, peuvent les simuler. Accéder au plan comportemental de nos interactions revient donc à découvrir la nature algorithmique du réel. »

Avec son équipe du MIT Media Laboratory, Pentland développe entre autres un système de vision artificielle pour modéliser – c'est-à-dire prédire – les interactions humaines dans des tâches de surveillance (Oliver *et al.*, 2000). Des agents comportementaux artificiels miment les comportements humains – décomposés ici dans des variables comme la vitesse de marche, la trajectoire, la proximité, l'évitement ou l'interaction avec d'autres agents, etc. – dans un espace virtuel. L'objectif est de pouvoir détecter un comportement menaçant en temps réel à partir d'un apprentissage supervisé où l'algorithme est dressé à reconnaître certains types de comportements. Le *pattern* d'un comportement menaçant que l'algorithme va reconnaître est un vecteur de certaines valeurs des variables. Les comportements humains sont simulés par des agents « synthétiques » qui produisent des données « synthétiques » qui, selon une logique bayésienne, deviennent le nouvel état *a priori* du modèle. « Le comportement de l'agent est déterminé par l'information contextuelle perçue : position actuelle, position relative à l'autre agent, vitesses, trajectoires, directions de la marche, etc., ainsi que par son propre répertoire de comportements possibles et d'événements déclencheurs » (Oliver *et al.*, 2000, p. 836, notre traduction).

La détection d'une anomalie ou la reconnaissance d'un comportement menaçant passe par sa prédiction. Tout comme un canon antiaérien qui tire le long de la trajectoire déjà parcourue par l'avion ne frappera jamais celui-ci, de même une caméra de surveillance ne pourra jamais dire si un comportement est menaçant par sa seule observation. Il faut, derrière l'écran qui rediffuse la scène surveillée, un agent humain ou artificiel qui interprète les comportements. Cette interprétation suit un protocole, mais un surveillant réellement efficace est celui qui a incorporé le protocole de surveillance au point où ce dernier constitue un système d'anticipation (au sens simondonien d'une



formalisation participant à un cycle de l'image ou ce que Bourdieu appellerait le « sens pratique ») préparant à de nouvelles expériences perceptivo-motrices. Dans la cour de récréation d'une école, un surveillant dégourdi n'attend pas la blessure provoquée par un jeu ayant perdu sa dimension ludique, il intervient avant que la blessure ne survienne et ramène éventuellement les comportements à l'intérieur de leur régime d'activité « normal ». Un agent alerte est un agent dont tout le schème perceptivo-moteur est amorcé, apprêté à recevoir une information incidente significative d'une menace, information qui va catalyser la suite de l'action amorcée par l'état d'alerte. Dans des cas extrêmes, le schème perceptivo-moteur excessivement alerté et amorcé est susceptible d'interpréter *n'importe* quel signal comme information (ce que Domingos appelle une fonction « hallucinée ») et risque de basculer dans un régime de quasi pure réactivité incontrôlable. Dans ces cas, le cycle de l'image est violemment rompu et l'anticipation, l'expérience et la formalisation, au lieu d'être des phases impliquant une invention, ne font que s'itérer à l'identique et s'hypertrophient dans un même amas indifférencié sous forme de stéréotype<sup>13</sup>. On reconnaît une « menace » dans la mesure où elle résonne anticipativement avec nos comportements amorcés par un régime d'information. L'anticipation est affective et transindividuelle. Si la cybernétique nous a appris que l'action et la prédiction sont toujours solidaires, Simondon nous permet d'insister sur le fait que le dynamisme de cette relation suppose toujours *l'expérience d'un problème* à partir de laquelle l'action prend une certaine valeur et à partir de laquelle une nouvelle norme est inventée.

En simulant les comportements humains, l'agent synthétique du système de vision artificielle prend activement part à la structuration des schèmes perceptivo-moteurs des acteurs impliqués dans l'activité surveillée. *Supervisé*, l'agent synthétique a été *dressé* à reconnaître certains comportements inédits plutôt que d'autres comme menaçants. Une des difficultés majeures consiste à ne pas classer par défaut comme menaçant un comportement réellement inédit ou très rare qui ne correspondrait pas aux paramètres du modèle et qui

---

13. Il pourrait être intéressant d'interpréter les actes de violence commis par la police, notamment les meurtres commis par des policiers sur des personnes *perçues* comme menaçantes, comme le résultat de schèmes perceptivo-moteurs structurellement prédisposés à amplifier et à sur-interpréter, selon un régime d'information stéréotypé, le comportement « menaçant » au point où ce dernier est perçu comme *déjà effectivement* violent. Il conviendrait, en outre, de directement inclure dans cette analyse les médiations techniques et leur participation à l'amplification ou la régulation des schèmes perceptivo-moteurs.

représenterait une anomalie<sup>14</sup>. Tout comme un parent pourrait dire à son enfant « reste près de moi, cette personne se comporte bizarrement » ou « il faut éviter des gens comme ça ». La différence tient évidemment au fait que là où l'énonciation du « comme ça » ou du « bizarrement » ne contient pas en elle toutes les conditions d'interprétation à partir desquelles une future occurrence pourra être dite équivalente, le modèle paramétrique sur lequel l'algorithme a été entraîné est censé établir un espace d'équivalences statistiques.

Sauf que dans le cas du *machine learning* il faut bien admettre que ce partage n'est sans doute plus si univoque et que les performances de la machine ont plus à voir avec les dominances normatives de l'activité dans laquelle elle participe que de ses potentialités ou limites inhérentes. Dans les travaux de l'équipe de Pentland, les données d'entraînement sont labellisées. Il existe donc un modèle statistique du comportement, ou un ensemble de comportements menaçants. On sait plus ou moins à l'avance quel type de comportement on veut que l'algorithme reconnaisse comme menaçant. Bien que les cas de *faux positifs* soient souvent présentés comme des dérapages, le fait est qu'ils traduisent simplement les stéréotypes raciaux, de sexe ou de classe qui se cachent comme autant de variables enfouies dans les données labellisées « menaçant ». Il ne faudrait pas s'étonner que l'« apprentissage » de l'algorithme confirme ces stéréotypes structurels. Qu'en serait-il, par contre, si au lieu d'entraîner l'algorithme sur un ensemble de variables constitutives de la menace on lui montrait simplement un ensemble de flux dans lesquels on estime qu'il y a un comportement menaçant sans lui dire de quoi il s'agit ? Est-ce que l'algorithme trouverait les mêmes régularités comportementales que nous pensions y voir ? Si l'on changeait le cadre d'activité, on pourrait également s'attendre à voir une modification des valeurs attribuées aux comportements. Forcément, ce scénario est invraisemblable dans la mesure où les algorithmes sont utilisés dans une activité de surveillance précisément pour surveiller, c'est-à-dire pour prédire une menace sur base d'une « probabilité réelle », d'une propension<sup>15</sup>. Hansen insiste sur le fait que les systèmes de prédiction reposent sur une forclusion du champ de données potentiel en délimitant un jeu de données actuel à partir duquel la prédiction pourra établir un calcul probabiliste fiable (Hansen, 2015, p. 126). Hansen rappelle qu'une prédiction suppose, tout en le circonscrivant, un champ de « potentialité réelle »

14. Olivier *et al.* (2000, p. 834). Sur la différence entre *anomal* et *anormal*, voir Canguilhem (1996, pp. 76-95) et pour une application aux techniques de détection, voir Pasquinelli (2015).

15. Pour une discussion de la « probabilité réelle », voir Hansen (2015).

ou de « surplus de sensibilité » : « Parce qu'elles comportent des données qui débordent sur celles qu'un système prédictif pourrait inclure, les données de l'efficacité causale du monde – les données qui constituent sa potentialité réelle – sont toujours en principe en train de faciliter des formes de savoir qui sont irréductibles à un programme particulier » (Hansen, 2015, p. 126, notre traduction). Le monde des données est toujours en excès sur son actualisation, il recèle toujours en son sein une charge d'indétermination qui subsiste aux tentatives de désamorçage par le système prédictif. De ce point de vue, on dira qu'*un algorithme est apprenant à partir du moment où son fonctionnement aura pour fonction d'actualiser cet excès de sensibilité, cette charge d'indétermination que recèle le champ de données* et dont la forclusion permet l'efficacité d'un système de prédiction (Hansen, 2015). Ce déplacement nous ramène, en apparence, à une sorte d'inductionnisme, qui pourrait, à bien des égards, nous rapprocher du fétiche qui voit dans la production et le traitement de « toujours plus de données » une panacée sociale. Il faut toutefois rappeler que le déplacement entrepris ici cherche justement à montrer qu'il n'y a rien d'absolument immanent à cette activité d'induction machinique, rien qui puisse garantir que l'excès de sens qu'elle présente soit aussi un excès d'objectivité ou de réalité, et ce précisément parce qu'elle est une *activité* dont les comportements se déploient dans des contextes sociaux spécifiques qui n'attendent pas qu'un sens émerge magiquement des données pour être formalisés et structurés. En ce sens, si inductionnisme il y a, c'est uniquement dans la mesure où il se déploie à l'intérieur d'un ensemble de déterminations sociales qui structure le champ des possibles, bien que ce champ soit sans cesse remis à jour et transformé par l'activité même d'apprentissage<sup>16</sup>.

## LE DEVENIR-IMAGE DE LA PRÉDICTION ALGORITHMIQUE

En reprenant ce geste spéculatif, il serait alors possible d'envisager le *machine learning* à travers le *devenir-image* des classifications qu'il opère. *L'image*

---

16. On retrouve, sous différentes formes, ce principe selon lequel la fin du développement phylogénétique est sans cesse reprise au début et transformée au long de l'ontogenèse individuelle, chez Simondon (2008) et Bourdieu (1997) aussi bien que chez Lev Vygotsky (1978, pp. 84-91), pour qui l'activité d'apprentissage est ce qui devance socialement le développement et non pas ce qui réalise une adaptation de l'individu à son environnement. Son concept de « la zone proximale de développement » permet de tenir compte de la différence entre le « développement actuel » de l'enfant et son « développement potentiel » induit par la présence directrice d'un adulte ou d'un autre enfant à un stade de développement plus avancé.

*non plus comme objet ou comme représentation, mais comme anticipation active de l'action.* Il paraît de plus en plus indiscutable que la prédiction, prise aussi bien comme formalisation comportementale et comme formalisation statistique des comportements, est un régime d'information auquel nous participons en permanence. Elle fait partie intégrante de nos comportements dans la mesure où ceux-ci participent toujours activement à une activité. L'activité normative est le *bruit* de la société. Bruit auquel nous livrons nos plus puissants algorithmes, dans l'espoir qu'ils y découvrent les formes cachées de nos habitudes, de nos manières de faire, l'inconscient social de nos gestes ; qu'ils anticipent ce que nous sommes incapables d'anticiper nous-mêmes. En somme, dans la rhétorique des tenants du *big data*, le *machine learning* pourrait être caractérisé comme une vaste et hétérogène entreprise destinée à forclure la possibilité même d'une *hystérésis*, de rendre impossible l'expérience d'un écart entre l'activité normative – entendue comme cycle de l'image ou comme incorporation de dispositions – et l'efficacité de la norme sociale. C'est ce qui fait dire à Cardon que « [l]e comportementalisme algorithmique est ce qui reste de l'*habitus* lorsqu'on fait disparaître les structures sociales » (Cardon, 2015, p. 71). Si cette formule permet de traduire de manière efficace un certain régime discursif largement répandu parmi les tenants du *big data*, lesquels espèrent accéder, grâce aux algorithmes du *machine learning*, à la couche comportementale la plus naturelle de nos comportements (en ce qu'ils ne seraient plus dénaturés par les effets des structures sociales), elle reproduit néanmoins ce que nous cherchions précisément à éviter : la fixation sur les discours au double sens où, d'une part, on estime ce dont est capable un algorithme à partir du discours qu'on tient à son égard et, d'autre part, on évalue les effets normatifs de l'algorithme lui-même comme une production discursive. À ce titre, le recours au terme « comportementalisme » est révélateur du fait qu'il est avant toute question de *discours sur* les comportements ou sur le comportement en ce qu'il est une stratégie discursive. *A contrario*, notre approche, sans chercher à minimiser l'importance des discours en tant que régimes d'énonciation à partir desquels la réalité d'un phénomène se construit, cherche davantage à prendre le comportement algorithmique lui-même comme enjeu d'une nouvelle individuation de la connaissance. Dès lors, il nous semble que la formule devrait plutôt s'énoncer de la manière suivante : « Les comportements d'apprentissage algorithmique sont ce qu'une machine fait de l'*habitus*. »

Dans une perspective simondonienne, nous avons proposé de concevoir le dynamisme du *machine learning* à partir du cycle de l'image. D'un point de vue sociologique, cela implique qu'il devrait être possible d'opérer une

projection statistique des comportements des algorithmes apprenants au même titre que ceux d'agents humains et d'interpréter leur distribution dans l'espace selon certains facteurs. Ce faisant, on pourrait montrer qu'il n'y a en réalité rien de purement automatique à cet apprentissage. Autrement dit, il serait possible non seulement d'utiliser les algorithmes pour « apprendre » de nouvelles corrélations à partir des comportements quantifiés des individus au sein d'une activité, mais plus encore, et ce serait un geste d'une réflexivité redoutable, d'étudier les comportements de ces mêmes algorithmes comme participant et informant l'activité dont ils sont censés extraire des données pour instituer des frontières, établir des corrélations, susciter des normes<sup>17</sup>. Aussi la constitution d'un savoir sociologique sur les comportements d'apprentissage algorithmique serait-elle l'occasion de réintroduire un écart discursif avec ce qui pourrait par ailleurs ressembler à une trop grande accointance avec les comportements pris comme seule expression du réel.

Il est possible, en suivant Bourdieu, d'éviter le subjectivisme sans tomber dans un objectivisme qui ferait des agents de simples automates sociaux, régis par les normes sociales comme les machines le seraient par les lois de la physique. Inversement, il devrait désormais être possible de soutenir que la machine elle-même n'est plus (mais l'a-t-elle jamais été) l'automate rêvé, l'exécutant docile, l'esclave parfaitement normé. En cherchant à se distancier d'une conception de l'agent social qui, même lorsqu'il obéit à la norme, représente, *bon gré mal gré*, une position souveraine grâce à laquelle c'est *lui*, en dernier ressort, qui décide d'obéir, Bourdieu a pu écrire qu'« [à] la différence de la commande, action sur une machine ou un automate qui opère par des voies mécaniques justiciables d'une analyse physique, l'ordre ne devient efficient que par l'intermédiaire de celui qui l'exécute ; ce qui ne signifie pas qu'il suppose nécessairement de la part de l'exécutant un choix conscient et délibéré, impliquant par exemple la possibilité de la désobéissance » (Bourdieu, 1997, p. 243). Autrement dit, la normalisation requiert l'activité normative. Activité qui est précisément ce qui fait une vie et ce qui manque à la machine. Or le *machine learning* nous confronte à la possibilité – dont il ne s'agit pas ici de

---

17. Soulignons que ce geste est déjà partiellement envisagé, sans qu'il ne soit porté au point d'une objectivation sociologique, par certains travaux dans le champ du *machine learning* qui cherchent à développer des techniques statistiques intégrées qui permettraient d'analyser le comportement de l'algorithme en train d'apprendre. Voir notamment Grimm *et al.* (2017). Voir aussi la proposition émergente, soutenue par certains chercheurs, de constituer une science interdisciplinaire du comportement machinique : <http://www.internetactu.net/2018/04/24/demain-des-psychologues-pour-robots/>, consulté le 28 avril 2018.

statuer sur la réalité ontologique ou la désirabilité éthique – qu’une machine apprenne à se comporter. Davantage qu’elle ne signale le caractère « algorithmique » de nos propres comportements, cette possibilité signifierait le fait que certains systèmes algorithmiques produisent des effets que l’on ne peut indexer à autre chose qu’à une activité normative. Les machines ne doivent pas être « conscientes » ou « intelligentes » pour participer aux cycles des images qui informent nos devenirs collectifs.

---

 RÉFÉRENCES
 

---

- BACHIMONT B. (2010), *Le sens de la technique : le numérique et le calcul*, Paris, Les Belles Lettres.
- BOURDIEU P. (1997), *Médiations pascaliennes*, Paris, Seuil, 318 p.
- BOURDIEU P. (2000), *Esquisse d'une théorie de la pratique*, Paris, Seuil, 448 p.
- CANGUILHEM G. (1996), *Le normal et le pathologique*, Paris, Presses universitaires de France.
- CARDON D. (2015), *À quoi rêvent les algorithmes : nos vies à l'heure des big data*, Paris, Seuil, 112 p.
- CHRISTIAN B., GRIFFITHS T. (2016), *Algorithms to Live By: The Computer Science of Human Decisions*, New York, Holt, 368 p.
- CLARK A. (2016), *Surfing Uncertainty: Prediction, Action and the Embodied Mind*, Oxford, Oxford University Press, 424 p.
- COLLINS H. (1990), *Artificial Experts: Social Knowledge and Intelligent Machines*, Cambridge MA, MIT Press.
- DESROSIÈRES A. (2008), *Pour une sociologie historique de la quantification*, Paris, Presses de l'École des Mines, pp. 10-11.
- DESROSIÈRES A. (2010), *La politique des grands nombres*, Paris, La Découverte.
- DESROSIÈRES A. (2014), *Prouver et gouverner*, Paris, La Découverte.
- DOMINGOS P. (2015a), *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, London, Allen Lane.
- DOMINGOS P. (2015b), *A Few Useful Things to Know about Machine Learning*, <https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf>, consulté le 24 mai 2015.
- FLACH P. (2012), *Machine Learning*, New York, Cambridge University Press.
- CANGUILHEM G. (1996), *Le normal et le pathologique*, Paris, Presses universitaires de France.
- GITELMAN L. (ed.) (2013), *'Raw Data' is an Oxymoron*, Cambridge MA, The MIT Press.
- GRIMM C., ARUMUGAM D., KARAMCHETI S., ABEL D., WONG L.-S., LITTMAN M. (2017), *Latent Attention Networks*, 31st Conference on Neural Information Processing Systems 2017.
- HANSEN M. (2015), « Our Predictive Condition; or, Prediction in the Wild », in Richard GURSIN (ed.), *The Nonhuman Turn*, Minneapolis, University of Minnesota Press, pp. 118-122.

- MACKENZIE A. (2015), « The production of prediction: What does machine learning want? », *European Journal of Cultural Studies*, vol. 18, n° 4-5, pp. 429-445.
- MALABOU C. (2017), *Métamorphoses de l'intelligence*, Paris, Presses universitaires de France.
- MARSLAND S. (2009), *Machine Learning: An Algorithmic Perspective*, Boca Raton FL, CRC Press, 406 p.
- OLIVER N. M., ROSARIO B., PENTLAND A. (2000), « A Bayesian Vision System for Modeling Human Interactions », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, n° 8, pp. 831-843.
- PASQUINELLI M. (2015), *Anomaly Detection: The Mathematization of the Abnormal in the Metadata Society*, Transmediale 2015, <http://matteopasquinelli.com/anomaly-detection/>, consulté le 13 janvier 2016.
- PENTLAND A. (2010), *Honest Signals, How they Shape Our World*, Cambridge MA, The MIT Press.
- REIGELUTH T. (2018), *Comporter la norme. La normativité de l'apprentissage algorithmique à partir du problème du comportement*, Thèse de doctorat, Bruxelles, Université Libre de Bruxelles.
- ROSENBLUETH A., BIGELOW J., WIENER N. (1943), « Behavior, Purpose and Teleology », *Philosophy of Science*, vol. 10, n° 1, pp. 18-24.
- ROUVROY A., BERNS T. (2013), « Gouvernamentalité algorithmique et perspectives d'émancipation », *Réseaux*, vol. 1, n° 177, pp. 166-167.
- RUYER R. (1954), *La cybernétique et l'origine de l'information*, Paris, Flammarion.
- SIMONDON G. (2008), *Imagination et invention (1965-1966)*, Chatou, Éditions de la Transparence.
- SIMONDON G. (2012), *Du mode d'existence des objets techniques*, Paris, Aubier, 367 p.
- SKINNER B. (1971), *Beyond Freedom and Dignity*, New York, Pelican Book, 216 p.
- VYGOTSKY L. (1978), *Mind in Society, The Development of Higher Psychological Processes*, Cambridge MA, Harvard University Press, 176 p.
- WIENER N. (1961), *Cybernetics or: Control and Communication in the Animal and the Machine*, MIT Press, Cambridge, MA, 212 p.
- WITTEN I., FRANK E. (2011), *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufman, Burlington, MA.