



Numérique ? Algorithme ? Big data ? Comprendre les technologies autour de l'Intelligence Artificielle

Sophie Agulhon

► To cite this version:

Sophie Agulhon. Numérique ? Algorithme ? Big data ? Comprendre les technologies autour de l'Intelligence Artificielle. Françoise Guénot. L'IA éducative. L'intelligence artificielle dans l'enseignement supérieur, Bréal, 2023, 2749555302. <hal-04871046>

HAL Id: hal-04871046

<https://hal.science/hal-04871046v1>

Submitted on 7 Jan 2025

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.



HAL Authorization

Numérique ? Algorithme ? Big data ?

Comprendre les technologies autour de l'Intelligence Artificielle.

Sophie AGULHON, Maître de Conférences, Université Paris 8 Vincennes-Saint-Denis/LED.

L'intelligence artificielle participe à une véritable révolution technologique qui bouleverse notre société. Elle fait partie d'une véritable grappe d'innovations, à l'origine des cycles économiques (Schumpeter, 1939) que nous connaissons depuis plus de trente ans :

- Naissance du World Wide Web et naissance des géants tels que Microsoft, Google ou encore Amazon dans les années 1990, suivi de l'explosion de la bulle Internet entre la fin des années 1990 et le début des années 2000,
- Apparition des réseaux sociaux, des monnaies basées sur la *blockchain* comme le bitcoin et la massification du recours aux applications pour *smartphones* dans les années 2000 et 2010,
- Renforcement de la gestion des données, développement accru de l'industrie 4.0, de l'Internet des objets (IoT), du Big Data et de l'Intelligence Artificielle depuis une dizaine d'année, etc.

La richesse de cette révolution et la prolifération technologique associée aux modèles économiques actuels et futurs nécessitent une mise à plat de plusieurs grandes notions dont l'enseignement supérieur doit se saisir pour assurer sa mission de formation comme de recherche.

1. Numérique, digital et système d'information

Numérique et digital (qui est un anglicisme) peuvent être considérés comme synonymes. Originellement, dans les domaines de l'informatique et des télécommunications, les deux sont des représentations d'informations ou de grandeurs physiques au moyen de caractères (tels que des chiffres), ou de signaux à valeurs discrètes. Au sein de ces représentations d'information, on retrouve les données numériques. Ces dernières sont des informations mesurables sous forme de nombres et structurées via une logique mathématique plutôt qu'en langage naturel. À cet égard, un ordinateur fonctionne en mode binaire 0 ou 1 bit. Historiquement, cela signifiait l'absence ou la présence d'un signal électrique, ce qu'il faut, aujourd'hui nuancer avec le recours à l'électronique de puissance et ses dispositifs de conversion. Ainsi, les données concernant la scolarité des étudiants telles que le nombre d'inscrits, les résultats aux examens ou encore le taux de décrochage, permettent de représenter les filières les plus performantes en matière de réussite universitaire et de classer les établissements sur ces critères. C'est ce que nous dénommerons le classement A.

Cependant, les usages du terme digital, tendent spécifiquement à le rapprocher des usages par les utilisateurs, mais aussi de la conduite du changement au sein des organisations en « transition digitale ». Plus largement, cette dernière revêt un autre intérêt pour les

organisations : l'ouverture de leurs systèmes d'information pour aboutir à la monétisation de leurs données numériques (aussi appelées data), via des schémas de valeurs plus ou moins complexes sur lesquels nous reviendrons.

Le système d'information se compose d'éléments physiques (par exemple, des individus et des équipements) et logiques (comme des processus ou encore des données numériques) organisés en réseaux de communication. Ils fonctionnent selon un jeu de procédures et de règles qui définissent comment collecter, stocker, traiter, communiquer ou retenir l'information, en fonction des besoins de l'organisation. Les organisations mettent, ainsi, en place un système d'information dont le rôle est d'assurer le bon fonctionnement de divers processus au sein de l'organisation et l'intégrité des données qui y sont associées. Sur lesquelles se bâtissent lesdits processus et l'information qui fait sens pour une organisation donnée (Reix, Fallery, Kalika et Rowe, 2016). En effet, afin d'être efficient, il est nécessaire d'associer des informations pertinentes, répondant aux besoins de certains acteurs et à un sujet particulier. En effet, toutes les données numériques ne sont pas toutes utiles au même niveau et pour tous les acteurs, c'est pourquoi l'information est à comprendre comme un résultat construit à partir de données choisies pour leur pertinence par rapport à une question particulière.

2. Algorithme et technologies de l'information

Si nous reprenons l'exemple du classement d'établissements de l'enseignement supérieur sur la base des trois critères évoqués plus haut (classement A), à savoir le nombre d'inscrits, les résultats aux examens ou encore le taux de décrochage, plusieurs critiques peuvent survenir :

1. Les étudiants inscrits au sein d'une formation le sont-ils par choix ou par absence de choix ?
2. Les étudiants de telle formation du supérieur disposent-ils d'une formation de l'enseignement secondaire en phase avec leur poursuite d'études ?
3. L'étudiant est-il originaire d'un milieu socio-économique favorable aux études : soutien financier, sensibilité à la question des diplômes, capacités de conseil et d'aide à l'orientation, etc. ?

Si nous prenons le 3^e point lié au milieu d'origine de l'étudiant, certains classements d'établissement intègrent des critères supplémentaires en prenant en compte, par exemple, l'adresse des étudiants, (avec des coordonnées GPS) donc des nombres, et en la croisant avec le revenu médian et la démographie du quartier concerné, par exemple.

Dans ce cas, le nouvel algorithme de classement pondère donc le nombre d'inscrits, les résultats aux examens, le taux de décrochage avec les données socio-démographiques et socio-économiques des adresses des jeunes. Néanmoins, plus les données socio-économiques et socio-démographiques recoupées avec l'adresse sont peu favorables aux études (personnes à faibles revenus, peu diplômées, etc.) alors que les résultats aux examens restent dans la moyenne, tout comme le taux de décrochage, plus il sera alors considéré que

l'établissement du supérieur concerné est investi dans la réussite universitaire de ses étudiants. Ce facteur modifie, in fine, le classement en considérant que ce qui pourrait constituer des pénalités à la poursuite d'études a été compensé. Ainsi, un algorithme se définit comme une suite d'instructions et d'opérations finie et non ambiguë, conçue pour résoudre une classe de problème (tri, association, transmission, etc.). Lorsque l'algorithme intègre une commande, il exécute cette dernière et facilite le traitement en masse de problèmes. Ainsi, enrichir l'algorithme de classement A consiste à affecter une sorte de bonus aux établissements dont le public est plutôt en difficulté socio-économique, mais dont les résultats aux examens sont finalement assez corrects pour les remonter dans ledit classement. Cela donnera naissance à un classement B.

3. Création de données et Big Data

Comme évoqué précédemment, l'une des dimensions de la transition digitale recherchée par les organisations est la mise en commun et l'exploitation des données détenues par les différents systèmes d'information de l'entité, mais aussi de celles détenues par d'autres parties prenantes et à l'externe (fournisseurs, clients, réseaux sociaux, etc.) pour créer de l'information utile pour la prise de décision.

Le Big Data désigne l'explosion du nombre et des types de données (Volume, Vitesse et Variété des données) depuis le début de l'informatique personnelle jusqu'à l'omniprésence du Web. Pour que la présence de volumes importants de données donne lieu à de la connaissance via l'information qu'elle contient, un champ de recherche extrêmement large et originellement indépendant de celui de l'intelligence artificielle s'est donc ouvert. Néanmoins, des convergences techniques sont à l'étude pour aboutir à une prise de décision intelligente.

Dans le champ de l'enseignement supérieur, le Big Data permettrait d'analyser les données en mouvement, d'analyser une grande variété et un volumes conséquent de données, ou encore de découvrir et d'expérimenter le potentiel d'une idée (Lemberger, Batty, Morel et Raffaëlli, 2015). Ainsi, le Big Data permet de prendre en compte tous les profils d'apprentissage qui peuvent se présenter, du plus commun au plus atypique et exceptionnel et de leur proposer, par exemple, un apprentissage adaptatif pour faciliter la réussite, en tenant compte des besoins de chacun. Par ailleurs, la diversité de données pouvant être combinées permet l'évaluation et le contrôle exhaustif des étudiants, mais aussi des prestations d'enseignement et de l'organisation administrative d'un établissement donné. À ce titre, l'exaspération autour d'un sur-usage des données et autour des mises en relations plus ou moins éthiques entre ces données via SQL (langage de communication avec des bases de données) pourraient susciter de vives tensions.

4. Machine Learning, Deep Learning et Intelligence Artificielle

Enfin, l'intelligence artificielle constitue à elle seule un champ scientifique particulièrement étendu, incluant le machine learning et le deep learning. L'intelligence artificielle inclut l'ensemble des techniques qui permettent aux ordinateurs d'imiter l'intelligence humaine. Dans ce cadre, le machine learning est un sous-ensemble qui permet aux machines, par apprentissage, de tirer des enseignements de leurs expériences pour améliorer l'exécution de leurs tâches. Selon Microsoft, ce processus d'apprentissage comprend plusieurs étapes : (1) alimenter un algorithme en données, (2) utiliser ces données pour entraîner un modèle, (3) tester et déployer le modèle, (4) utiliser le modèle déployé pour effectuer une tâche prédictive automatisée¹.

Enfin, le deep learning fonctionne au moyen de réseaux neuronaux artificiels qui comprennent plusieurs couches qui s'alimentent pour démultiplier le processus d'apprentissage au travers de son propre système de traitement de données. Autrement dit : si une couche apprend sur la requête X pour corriger sa prochaine requête Y, une autre couche peut aussi apprendre des résultats de la requête X et améliorer ses propres fonctions, et qui peuvent être de nature tout à fait différentes de la requête X. C'est ce qui explique notamment la puissance de ChatGPT et le fait que cette technologie questionne tant l'enseignement supérieur aujourd'hui.

En conclusion, la Data Science comme horizon ?

Les schémas créateurs de valeur par la digitalisation peuvent s'étendre de la pratique digitale ~~et~~ à l'activation de la donnée, jusqu'à la création de nouveaux écosystèmes, de nouveaux modèles de prix et de coûts, en dehors de la chaîne de valeur traditionnelle (Zeitoun, 2022). Dans ce contexte, il est important que l'enseignement supérieur soit au fait des nuances des modèles qui se développent au sein des organisations et qui s'offrent aussi à lui pour créer du sens entre science, technique et société.

Références

Lemberger, P., Batty, M., Morel, M., Raffaëlli, J.L. (2015). *Big Data et Machine Learning. Manuel du data scientist*. Paris : Dunod.

Reix, R., Fallery, B., Kalika, M. & Rowe, F. (2016). Chapitre 1. La notion de système d'information. Dans : R. Reix, B. Fallery, M. Kalika & F. Rowe (Dir), *Systèmes d'information et management* (pp. 1-69). Paris : Vuibert.

Schumpeter, J. (~~1982~~) (1939). *Business Cycles*, vol. 1 et 2. Philadelphia : Porcupine Press.

Zeitoun, H. (2022). *Datae Humanum. La transformation digitale au service du renouveau humain*. Paris : La Sirène aux Yeux Verts.

<https://www.oracle.com/fr/artificial-intelligence/deep-learning-machine-learning-intelligence-artificielle.html>

¹ <https://learn.microsoft.com/fr-fr/azure/machine-learning/concept-deep-learning-vs-machine-learning>

