



# DE LA BI À L'IA

## UN CHANGEMENT DE PARADIGME DANS LA CRÉATION DE VALEUR À PARTIR DE LA DONNÉE

**David TSANG HIN SUN**  
Expert Data Science & IA | Keyrus



La Business Intelligence a été au cœur des investissements des entreprises il y a deux décennies. Sa démocratisation a été un tournant dans la manière de mobiliser les données pour améliorer le pilotage de l'activité et les décisions opérationnelles. Aujourd'hui l'accès à l'intelligence artificielle permet aux entreprises de franchir une nouvelle étape : se réinventer en créant de la valeur à partir de données jusque-là sous-exploitées ou inexploitées.

Les avancées scientifiques et techniques de ces 10 à 15 dernières années ont déplacé les projecteurs longtemps

braqués sur l'informatique classique - traitant et produisant des données structurées - vers l'intelligence artificielle (IA). Il faut voir cette dernière non pas comme une solution "magique" à toutes les problématiques, mais comme **un ensemble de méthodes, de techniques et d'outils** qui repoussent les limites auxquelles se heurtait le traitement des données et, par conséquent, la valeur que les entreprises pouvaient espérer en tirer. Mais si les données sont bien le carburant de l'IA, comme elles étaient déjà celui de la Business Intelligence (BI), il est essentiel de comprendre le changement d'échelle et de nature que ce carburant a connu au début du XXI<sup>e</sup> siècle avec **l'émergence des Big Data**, impossibles à exploiter avec les méthodes et l'outillage préexistants.

## BIG DATA, BIEN PLUS QU'UNE RUPTURE QUANTITATIVE

Les Big Data sont la résultante de plusieurs évolutions concomitantes : la montée en puissance d'internet et du web ; la bascule vers le numérique des anciens canaux de communication et des médias analogiques ; l'informatisation générale des processus tertiaires et l'accélération de l'automatisation industrielle. Avec la multiplication des systèmes, leur diversification et l'intensification des usages numériques, **les données ont évolué selon trois dimensions, les fameux « 3V » - volume, vitesse, variété** - communément utilisés pour caractériser les Big Data :

**La volumétrie** des données augmente de façon exponentielle, ce qui dans un premier temps pose aux entreprises des problèmes de stockage et, très vite, un problème de capacité de traitement ;

**La vitesse** à laquelle sont produites les données s'accroît elle aussi, et est incompatible avec la plupart des outils d'analyse en place, conçus pour des traitements en mode batch et non pour **une exploitation en temps réel et au fil de l'eau** ;

**La variété** des formats de données est elle aussi un phénomène nouveau pour les entreprises qui, jusque là, ne traitaient que des données structurées et, dans une moindre mesure, semi-structurées. Or, l'activité génère de plus en plus de données non structurées - texte, son, image, vidéo - dont **la valeur potentielle est pressentie**, mais pour lesquelles les outils de traitement de la BI traditionnelle sont inopérants.

Les infrastructures et systèmes d'information classiques ne sont ni dimensionnés ni structurellement conçus pour prendre en charge les Big Data. Ce sont les acteurs majeurs de l'internet qui, les premiers, vont non seulement démontrer la richesse des Big Data en développant des usages inédits de la data, mais aussi ouvrir le jeu, en publiant au milieu des années 2000, **des méthodes permettant de manipuler et traiter les données massives**. On pense notamment à Google BigTable, Hadoop et MapReduce

qui, en rendant **le stockage et le calcul** distribués accessibles aux entreprises, marquent un virage dans l'exploitation de la data et sa valorisation. S'appuyant sur ces techniques, les entreprises commencent alors à créer des data lakes pour stocker de manière décloisonnée les données de toutes origines et de tous formats. L'outillage pour exploiter cette matière première n'est cependant pas encore complet.



## DEEP LEARNING, UN BOND EN AVANT DÉCISIF

Une nouvelle étape est franchie en 2012 avec l'apparition de méthodes scientifiques et de techniques particulièrement **adaptées à l'indexation et au traitement de l'image, du son et de la vidéo**. Ces catégories de données deviennent en effet prépondérantes avec l'essor des réseaux sociaux et de la vidéo *on demand*, des usages massifs tirés par la généralisation du haut débit et la démocratisation du smartphone.

En laboratoire, des avancées scientifiques telles que les réseaux de convolution, viennent étendre les capacités des algorithmes d'apprentissage statistique (Machine Learning) et, ce faisant, donnent un nouvel élan à la recherche en intelligence artificielle. Basé sur des réseaux de neurones artificiels, **le Deep Learning** (sous-ensemble du Machine Learning) a en réalité vu le jour dans les années 1980. Perfectionnés et couplés à l'utilisation des processeurs graphiques (GPU) autorisant des calculs distribués plus performants que les classiques CPU, les algorithmes de Deep Learning ouvrent la voie, à partir de 2012, à des applications inenvisageables auparavant – notamment en matière de **prédiction, de reconnaissance visuelle et de traitement du langage naturel**.

Deux facteurs supplémentaires vont permettre à ces techniques de sortir des laboratoires de recherche et de se diffuser rapidement à grande échelle :

**la démocratisation du cloud**, qui élimine la nécessité pour les entreprises d'investir en propre dans des infrastructures de stockage et de calcul distribués dimensionnées pour ces nouveaux types d'applications ;

**le libre accès aux bibliothèques d'algorithmes et aux bases d'apprentissage**, permettant aux développeurs de concevoir et configurer beaucoup plus rapidement des solutions répondant, d'une part, à des besoins insatisfaits par la BI et, d'autre part, à des cas d'usage inédits.

Ces facteurs et évolutions combinés expliquent le décollage du marché de l'IA. Devant la dynamique actuelle de la demande, Gartner estime que ce marché connaîtra **une croissance de 36% entre 2020 et 2024**<sup>1</sup>. Il est à noter que, sur la même période, les prévisions de croissance du marché de la BI sont de seulement 12%.

1. Source : Forrester Research / Gartner Research

## INTÉGRER L'IA AUJOURD'HUI – QUELLES APPLICATIONS POUR QUELS BÉNÉFICES ?

Si tout le monde a en tête des applications de l'IA très médiatisées, telles que la voiture autonome ou la lecture des images médicales, toutes les entreprises ont aujourd'hui la possibilité de s'approprier ces techniques, méthodes et outils pour :

**améliorer leurs performances opérationnelles**, en simplifiant et automatisant leurs processus internes et externes ;

**renforcer leur agilité et leur capacité à anticiper** grâce au pilotage par la donnée et à l'analyse prédictive ;

**se démarquer de leurs concurrents** en proposant des offres et des services innovants qui créent de la valeur pour leurs clients, partenaires et autres parties prenantes.

Pour illustrer le potentiel et la valeur ajoutée de cette rupture technologique, voici quatre projets, très différents, que les data scientists et experts en intelligence artificielle de **Keyrus** ont accompagné et contribué à mettre en œuvre chez des clients.



### GÉNÉRATION AUTOMATIQUE D'ALBUMS PHOTOS

Ce leader du tirage photo a voulu proposer à ses clients un service ultra-personnalisé s'inspirant de ce que fait Google en matière d'indexation d'images. À partir des photos chargées par le client sur le site, un algorithme de reconnaissance d'images classe les photos par thèmes (paysage, fête, animal de compagnie, etc.). Les photos ainsi classées sont ensuite organisées en albums thématiques dont le client peut commander l'impression en un simple clic. Effet « Waouh ! » pour le client et gain de temps (il n'a pas à composer lui-même son album). Cette innovation est également source de commandes supplémentaires pour l'entreprise sur son cœur de métier : l'impression photo.

### MAINTENANCE PRÉDICTIVE

Ce fournisseur de matériel d'élévation de charges veut maximiser le taux d'utilisation de son parc de machines et maîtriser ses coûts de maintenance en intervenant de manière préventive sur les machines susceptibles de tomber en panne ou de connaître des défaillances chez ses clients. L'analyse statistique des données du passé et l'analyse au fil de l'eau des données des capteurs connectés (IoT) dont sont équipées les machines (surveillance de la température, du niveau d'huile, de la pression...) permettent de détecter les signes avant-coureurs d'une panne et de déclencher une intervention chez le client à des fins de vérification, de réparation ou de remplacement des pièces à risque. L'entreprise évite ainsi des pannes potentiellement plus longues et plus coûteuses à réparer et renforce la fiabilité de ses machines, ainsi que la qualité globale du service rendu aux clients.

### DÉTECTION & PRÉVENTION DES RÉSILIATIONS

Ce grand fournisseur d'énergie cherche à identifier le plus en amont possible les clients susceptibles de changer de fournisseur afin d'engager des actions marketing ciblées pour les retenir. Un algorithme de Machine Learning, entraîné sur les données historiques d'une très large base de clients, a identifié dans ces données les invariants qui caractérisent les clients qui ont résilié leur contrat dans le passé. Grâce à cet apprentissage, l'algorithme est capable de prédire à partir des données actuelles, quels sont les clients d'aujourd'hui présentant ce risque à un niveau plus ou moins élevé. Les équipes marketing peuvent alors adresser à ces clients des offres et des communications personnalisées visant à les retenir et à les fidéliser.

## LA RÈGLE D'OR : PARTIR DU BESOIN, PAS DE LA TECHNOLOGIE

Il y a 10 ans, aucun de ces projets n'aurait pu voir le jour. S'ils ont pu être mis en œuvre, c'est certes parce que les technologies sont disponibles, mais surtout parce que les entreprises concernées n'ont pas cherché à « faire de l'IA pour faire de l'IA ». Elles ont au contraire cherché à répondre à un besoin clairement identifié, **sans préjuger des technologies à mobiliser**. Il se trouve que la réponse la plus pertinente et la plus efficace à ce besoin pouvait être apportée par une combinaison d'algorithmes et de techniques d'apprentissage automatique. Ce n'est pas toujours le cas ! En matière d'automatisation de tâches et de processus, par exemple, les solutions de RPA (Robotic Process Automation) peuvent satisfaire de très nombreux besoins et apporter une réelle valeur à l'entreprise. En revanche, pour les cas d'usage impliquant l'analyse et la classification de gros volume de données ou de données complexes (comme les images), le Machine Learning ou le Deep Learning est souvent incontournable.

De l'analyse des sentiments aux agents conversationnels intelligents, en passant par la reconnaissance faciale et la prédiction de pannes ou de comportements, les techniques d'apprentissage statistique, couplées aux technologies de computer vision, de traitement du langage naturel et à l'IoT, ouvrent des possibilités qui sont loin d'être épuisées. **Ces possibilités ne sont pas illimitées** pour autant, pour des raisons d'ordre scientifique, mais aussi pour **des raisons éthiques dont les entreprises doivent prendre la pleine mesure**. Pour intégrer ces avancées et en tirer durablement de la valeur, les entreprises doivent gagner en maturité sur ces sujets. La priorité pour elles, n'est pas d'internaliser toutes les compétences en data science, en mathématiques et en probabilités statistiques que requiert la maîtrise de ces techniques. Il est indispensable en revanche qu'elles prennent conscience des biais que peuvent comporter les algorithmes, de même que les données utilisées pour leur apprentissage, et qu'elles comprennent **les limites inhérentes à la nature statistique et probabiliste de cet apprentissage**.

En raison de ces limites, l'IA d'aujourd'hui ne peut pas traiter efficacement certains problèmes. Mais la recherche scientifique se poursuit et l'effervescence et la créativité qui animent actuellement le monde de la recherche se traduiront – demain, après-demain ou plus tard – par de nouvelles avancées qui contribueront à l'amélioration de nos vies. Au vu de l'orientation actuelle des investissements dans la recherche en IA, **des progrès sont à attendre du côté de l'apprentissage non supervisé**.

Dans ce domaine comme dans d'autres, quel que soit le rythme des percées scientifiques, la créativité humaine leur trouvera des applications concrètes pour résoudre de nouveaux problèmes et faire face à de nouveaux défis. Mais n'oublions jamais que pour être vraiment acceptées par la société et lui être utile, ces applications, comme les technologies sur lesquelles elles reposent, doivent **respecter des principes éthiques reconnus et acceptés par tous les acteurs**. C'est pourquoi – de OpenAI, présidée par Elon Musk, à la Déclaration de Montréal « pour un développement responsable de l'intelligence artificielle », portée par Yoshua Bengio, en passant par les *Lignes directrices en matière d'éthique pour une IA digne de confiance* de la Commission européenne – de nombreuses initiatives visant la mise en place d'un tel cadre éthique pour l'IA ont vu le jour et doivent être encouragées.

D.T.H.S.

## À PROPOS DE L'AUTEUR



### David TSANG HIN SUN

Après des études en mathématiques appliquées (Maîtrise de mathématiques - Calcul numérique), David a travaillé pendant plus de 15 ans dans la recherche en tant qu'ingénieur.

Passionné par la donnée, les Big Data, l'IA et le Machine Learning, il rejoint le secteur privé en 2015 autour de projets Big Data et d'Intelligence Artificielle à la SNCF.

C'est en novembre 2019 qu'il rejoint **Keyrus** pour développer et animer la communauté IA & Data Science.

## KEYRUS

### Keyrus, créateur de valeur à l'ère de la Data et du Digital

Acteur international du conseil et des technologies, spécialiste de la Data et du Digital, **Keyrus** a pour mission d'aider les entreprises à tirer profit du paradigme de la Donnée et du Numérique pour accroître leur performance, faciliter et accélérer leur transformation et générer de nouveaux leviers de croissance, et de compétitivité.

Plaçant l'innovation au cœur de sa stratégie, **Keyrus** développe une proposition de valeur unique sur le marché autour d'une offre novatrice qui s'appuie sur la combinaison de trois expertises majeures et convergentes :

#### Data Intelligence

Data Science - Intelligence Artificielle - Big Data & Cloud Analytics - Business Intelligence - EIM - CPM/EPM

#### Digital Experience

Innovation & Stratégie Digitale - Marketing Digital - DMP & CRM - Commerce Digital - Performance Digitale - User Experience

#### Conseil en Management & Transformation

Stratégie & Innovation - Transformation Digitale - Pilotage de la Performance - Accompagnement des Projets

Présent dans 20 pays et sur 4 continents, le Groupe Keyrus emploie 3 000 collaborateurs.

Plus d'informations sur [www.keyrus.fr](http://www.keyrus.fr)