

Le Big Data, mine d'or des entreprises, fléau de l'écologie

Introduction

Pour mon essai, j'ai décidé de choisir des thèmes qui me tiennent à cœur notamment l'écologie et sa protection. Il m'est important de changer mes habitudes et de faire évoluer celles de tous afin de ralentir le changement climatique et ses effets négatifs qui s'abattent sur notre planète. Ces bouleversements seront toujours de plus en plus fort si nous ne faisons rien. Le second thème de mon essai portera sur les avancées technologiques, spécifiquement le Big Data. Nous avons toujours favorisé les avancées technologiques au détriment de notre environnement surtout quand il s'agit de faire fructifier les ressources économiques d'une entreprise ou d'un pays. Nous vivons dans une économie capitaliste et dans un monde en perdition en raison des habitudes de consommation. Le sujet du Big Data, mine d'or des entreprises, fléau de l'écologie s'est donc présenté tout naturellement à moi.

Il m'était important de présenter un sujet éthique. Un sujet éthique est un sujet qui intègre une vision et des jugements moraux. Une vision morale rassemble les valeurs, les normes et les règles de conduite de notre société. Mon sujet est basé sur le respect de l'environnement, contre sa destruction massive ; il est donc pleinement éthique et fut à l'ordre du jour de la COP26 qui a eu lieu à Glasgow en Ecosse et qui a pris fin le 13 novembre 2021.

Les technologies liées à Internet, les Cloud, les intelligences artificielles ont réalisé une augmentation exponentielle des données produites. Aujourd'hui toutes nos informations personnelles, professionnelles, les informations sur les acteurs économiques et autres sont récoltées, stockées dans des milliers d'ordinateurs et processeurs ultrasophistiqués et puissants. Ces machines sont capables d'analyser les données en un temps record puis de les distribuer suivant les besoins. Le volume de ces données structurées ou non est telle qu'il a été nommé Big Data au sens littéral du terme « grande donnée » en raison de son volume. La masse d'information est telle qu'il est impossible de la traiter à l'aide des outils traditionnels tels que les bases de données ou les logiciels. Pour la majeure partie des entreprises, le Big data représente un tel volume de données, qu'il dépasse la capacité de traitement opérationnelle.

Où en est-on de l'écologie dans le monde ?

Les ressources ne sont pas immuables, elles n'ont pas la capacité de se renouveler de manière instantanée. Si nous prenons l'exemple d'un arbre nous mettons moins de temps à couper un arbre que celui-ci met de temps à pousser. Il en est de même pour la forêt ; nous consommons plus de bois que celui-ci met de temps à se régénérer. Les scientifiques ont donc calculé le jour dit du dépassement

« Earth Overshoot Day », date à laquelle nous avons consommé la totalité des ressources que la Terre peut renouveler en un an, en prenant en compte la déforestation, la surpêche, les émissions de gaz à effet de serre et la surexploitation agricole. Cette année, le jour du dépassement a eu lieu le 29 juillet. Certains événements tragiques peuvent être perçus différemment quand nous nous apercevons que le covid19 a fait reculer de trois semaines le jour du dépassement pour 2020. Nos mœurs et conditions de vie ont tellement évolué dans une surconsommation que depuis les années 2020 nous avons perdu 1 trimestre, nous sommes passés de fin septembre à fin juillet. Alors que depuis 1970 nous étions au 29 décembre, en 30 ans nous avons atteint la date du 23 septembre. L'accélération de la surconsommation de la terre est donc récente ; elle s'est faite dans ces vingt dernières années.

Depuis 1958 de nombreux scientifiques mettaient en lumière le début du changement climatique ; en 1970 on parlait d'un possible réchauffement climatique capable de faire fondre la banquise dans 200 ans en raison des émissions de CO₂. En 1988, le scientifique de la NASA James Hansen témoigne devant le Congrès d'un changement climatique bien réel. C'est seulement en 2015 que la majeure partie des pays du monde ont pris la décision et l'engagement de réduire leurs émissions de CO₂. Précédemment les plus gros pollueurs n'avaient pas fait le choix d'être présents. Pour certains présidents comme pour l'ancien président américain, Trump, faire évoluer la manière de consommer des individus afin de réduire les émissions de CO₂ ferait obligatoirement baisser la croissance du pays. Pour lui le choix est simple il serait soit de faire le choix en faveur de la croissance du pays soit de protéger le climat.

La Conférence des Parties 26 (COP26) permet à 120 dirigeants du monde entier de se réunir chaque année, le but étant de vérifier le respect de leurs engagements de l'année passée, de prendre de nouveaux objectifs et de mettre en place de nouvelles actions. Depuis 1995, les gouvernements se regroupent, chaque année, pour essayer de trouver ensemble des solutions mondiales à l'urgence climatique. En 2015, a eu lieu la COP21 à Paris ; chacun des pays se sont engagés à respecter leur contribution déterminée au niveau national (CDN). Chaque année, de nombreux absents restent à déplorer ; cette année encore des pays d'importance majeure n'ont pas été présents, comme la Chine ou la Russie. Les objectifs sont de taille si nous ne voulons pas voir notre planète comme nous la connaissons complètement détruite. Il est primordial de ne pas dépasser un réchauffement climatique de plus de 1,5°C, d'atteindre une neutralité carbone, de protéger les habitats naturels de toutes les destructions telles que les feux de forêt ou « mégafeux », la montée des eaux, les inondations, les cyclones.

En 1993, le niveau moyen des mers du globe n'avait pas augmenté alors qu'en 2020 les mers avaient augmenté de 8cm. Cela n'est pas dû aux hasards mais bien au réchauffement climatique qui fait fondre les calottes glaciaires. Depuis 2002 le Groenland a perdu 4 000 gigatonnes de glace, ce qui représente 11 gigatonnes par an. Dans 78 ans pour la fin du siècle d'après une étude menée par le GIEC la mer aura augmenté de plus d'un mètre si nous respectons le fait de ne pas dépasser une augmentation de 1,5 °C supplémentaire de la température de la terre il n'y aurait alors qu'un demi-milliard d'humains qui vivra dans une ville inondée mais si nous dépassons ce 1,5°C de 0,5°C alors ce sera 200 millions d'habitants en plus qui seront impactés quotidiennement. En 1990, nous étions au degré 0,40 ; en 1920 nous étions au degré 0, alors que cette année nous atteignons presque les 1,20 degré supplémentaire.

L'ère des mégafeux, laissant derrière eux des millions d'hectares de terre calcinés et sans vie, est déjà en marche. Malgré ce que nous pourrions penser, les feux de forêt ne sont pas uniquement présents dans des zones chaudes ils sont présents partout dans le monde jusque dans la toundra de Sibérie. EN 2020, plus de dix millions d'hectares de végétation sont partis en fumée, plus 1,25 milliards d'animaux ont été tués. L'Australie utilise plus des 2/3 de sa consommation en électricité par les énergies fossiles comme le charbon. La France rejette 4 tonnes de CO₂ par an quand il est de 15 tonnes pour l'Australie puisque le dirigeant du pays est climato-septique.

Quel rôle joue le Big Data pour les entreprises ; quel en est son intérêt ?

Chaque jour, l'entreprise doit prendre des décisions importantes en termes de stratégie, stratégie d'innovation, stratégie de distribution, stratégie de communication ou encore stratégie commerciale et ainsi dans une vision beaucoup plus opérationnelle, sur le terrain, pour développer ses forces commerciales afin de conserver ses parts de marché et les développer. Dans son développement l'entreprise a besoin d'être très bien informée sur son environnement, sur son marché, ses concurrents, ses clients, futurs prospects, la législation et la réglementation en vigueur ainsi que sur les attentes de ses parties prenantes.

Les données deviennent donc un des éléments le plus précieux pour toutes les entreprises des temps modernes. L'entreprise doit réaliser 3 étapes afin de pouvoir utiliser de manière pertinente l'information. Dans un premier temps, elle doit utiliser tous les leviers pour capter de l'information en masse (le Big Data), puis elle doit analyser l'information pour déterminer de la véracité et du bien-fondé de l'information ; s'assurer que l'information provient d'une source de confiance ; et enfin elle doit conserver l'information pour ne pas que celle-ci soit perdue mais reste disponible à la bonne personne, au moment opportun. La bonne information doit pouvoir être disponible au bon moment par la bonne personne et que celle-ci soit informée de la présence de cette information et de son emplacement afin qu'il puisse l'utiliser et que celle-ci devienne un avantage et un outil d'aide à la décision.

Pour travailler plus efficacement, l'entreprise ne peut pas se contenter d'utiliser des outils d'analyse traditionnelle pour analyser le Big Data car la quantité d'information et son trafic sont beaucoup trop importants. De plus, l'humain influe forcément d'une manière ou d'une autre sur les informations trouvées et risquerait de faire perdre ou d'omettre des informations capitales en fonction de ses propres connaissances et de sa vision du marché.

Le Big Data est récent, créé grâce à internet, du Cloud et des intelligences artificielles qui ont entraîné une véritable explosion du nombre de données produites à l'échelle de l'humanité. Il représente aujourd'hui environ 2,5 trillions d'octets de données créées tous les jours. Le Big Data analytique est donc la voie actuelle vers le profit.

Il existe trois types d'analyse : prescriptive, prédictive et descriptive. Le ciblage comportemental consiste à adapter la publicité en ligne à ce que l'internaute recherche. Les informations étudiées sur la base de données clients permettent de connaître avec précision les tendances et les modèles des clients afin de leur proposer une offre et des « recommandations produits » en fonction de son profil et intérêts. Ce qui permet d'augmenter la fidélisation des clients en observant leurs tendances, de ne commercialiser les produits qu'en ne conservant à l'esprit les attentes réelles des clients.

Aucun type d'information n'est laissé au hasard que ce soit sur les données démographiques, les données de localisation, le niveau de trafic d'une rue ou encore les données personnelles clients telles que le nom, l'adresse, l'historique de recherche et de paiement sont étudiés. Par exemple deux géants pour qui nos données sont cruciales, Amazon pour offrir une véritable expérience personnalisée et Starbucks pour choisir avec soin l'implantation de ses boutiques sur des zones extrêmement passantes et créant des boutiques proches les unes des autres pour réduire le risque.

Une autre facette du Big Data est la conservation des données. Le stockage des données est réalisé dans des *Data centers* (centre de données). Un *Data center* est un lieu dans lequel se trouvent une multitude d'ordinateurs, de baies de stockage et des multitudes de processeurs groupés en réseau dans des usines gigantesques. Ces espaces de stockage sont vendus aux entreprises afin qu'elles puissent disposer d'une immense infrastructure informatique pour stocker site internet, outils d'analyses, processeurs et outils informatiques ainsi que de stocker et de traiter des quantités importantes de données. L'achat d'espace dans un centre de données, permet aux entreprises de profiter du savoir et d'une technologie de pointe à moindre coût tout en réduisant leur dépense car elles n'ont pas besoin d'acheter leur machine, de verser des salaires, de construire un espace dédié prenant en considération les problèmes liés à la sécurité, la connectivité, aux besoins de connexion en très haut débits.

En quoi la Big Data impacte négativement l'écologie qu'elles solutions peuvent les résoudre ?

L'impact négatif de la Big Data

Comme nous l'avons vu précédemment, depuis le début 1990, la technologie n'a cessé d'évoluer en faveur des entreprises et de leurs utilisateurs ; la révolution technologique est donc un événement

récent. La technologie dans son évolution n'a pas pris en considération d'allier prouesse technologique et environnement.

Chaque machine a besoin d'éléments bien spécifiques pour sa construction et nous engendrons continuellement de nouvelles données sans forcément nous en rendre compte. Par exemple, lorsque nous enregistrons nos données sur un site, lorsque les sites nous envoient continuellement des mails ou sms par le biais de leurs newsletters ou encore tout simplement en acceptant la transmission de nos données et les cookies ou encore par le développement du *cloud computing* qui nous permet de stocker toutes nos données de manière dématérialisée afin d'y avoir accès de n'importe où et n'importe quand, simplement grâce à une connexion internet. Nous avons pu constater une explosion quantitative des données numériques produites à l'échelle de l'humanité. Toutes ces actions réalisées à titre personnel ou professionnel comme les bases de « données clientèle » sont des actions actives ou passives, énergivores et impliquent des exigences énergétiques de plus en plus élevées pour les centres de traitement et de stockage.

En 2012, il existait plus de 500 000 data centers et le nombre de ces derniers ne cesse d'augmenter ; l'avancement de la technologie actuelle telle que l'intelligence artificielle, les outils d'intelligence économique, la veille et bien d'autres outils engendrent continuellement de nouvelles données, ce qui implique des exigences énergétiques de plus en plus élevées. Les centres de traitement aujourd'hui dépassent les 8 millions.

Le volume de données générées par l'humanité devrait augmenter de 27% par an au fil des années à venir ce qui est énorme lorsque nous savons qu'elle représente 912,5 trillions d'octets de données créées tous les ans. Les conséquences vont être désastreuses et amplifier avec l'apparition de la 5G qui permettra à chacun d'entre nous de transférer de larges volumes de données à très haute vitesse, obligeant les prestataires à renouveler la totalité de leurs installations et les utilisateurs à racheter de nouveaux outils compatibles ou encore par l'accroissement de l'utilisation d'appareils IOT autrement dit l'internet des objets qui accroît la collecte des données de manière constante et sans fin. La consommation énergétique est déjà très importante et augmenter la puissance et la vitesse de transmission ne fera qu'accroître le besoin et la puissance énergétique nécessaire.

Les data centers ne sont pas en reste quant à leur consommation énergétique. Un data center est un lieu où est regroupé de nombreux ordinateurs allumés et travaillant sans discontinuer, en toutes circonstances. En plus de leur consommation énergétique en continu, les ordinateurs surchauffent. Il est obligatoire de prévoir des systèmes de refroidissement afin qu'ils ne saturent pas et puissent fonctionner sans interruption. La consommation énergétique des centres de données a augmenté entre 2017 et 2022 de 31% ; aujourd'hui elle représente 416,2 térawattheures. A titre de comparaison celle de la Grande-Bretagne est de 300 térawattheures et son empreinte carbone peut être comparée au secteur du transport aérien. Il faut savoir que 40% de l'énergie consommée par un data center est utilisée par le système de refroidissement.

Les installations ont accumulé nombre de mauvaises installations et de mauvaises pratiques notamment concernant le rejet de produits utilisés pour les systèmes de refroidissement, faisant appel à des substances nocives et chimiques ainsi qu'à des batteries de réserve, en cas de défaillance de l'alimentation électrique. Ces batteries possèdent des éléments toxiques après usage sans compter le fait qu'afin de répondre aux besoins de ces installations, il est nécessaire d'obtenir en masse des matériaux miniers rares.

Toutes ces activités liées au Big Data est extrêmement polluant et nuisible pour la Terre ; c'est un véritable barrage pour la protection de l'environnement et les prévisions dans cette surconsommation ne sont pas bonnes car en 2030 rien que les Data centers représenteront 10% de la production mondiale électrique et en 2040 ils représenteront ¼ contre seulement 3% actuellement.

Qu'elles solutions peuvent être trouvées ?

Avant de réfléchir à de nouvelles solutions, nous pourrions travailler sur une utilisation plus raisonnée du Big Data en convaincant les entreprises à ne stocker que les données primordiales et non les informations superficielles, en imposant une durée maximale de conservation de nos données, favorisant la simplification des demandes d'oublies.

Pour réduire et limiter la hausse de l'impact environnemental des Data centers Il faut ainsi chercher à les rendre plus respectueux de l'environnement. Dans un premier temps, en s'attaquant au problème de consommation des systèmes de refroidissement, en favorisant leur construction dans des zones au climat froid et par le biais d'une aération bien maîtrisée. Il ne serait pas nécessaire de passer par un système de climatisation qui, du fait de la transformation de l'air chaud en air froid, demande une surconsommation extrême d'énergie. Faire le choix d'installer ces centres dans une zone froide plus que dans une chaude est déjà une solution plus respectueuse de l'environnement de nombreux « Green Datacenter » voient le jour. Les Green Data centers font le choix de réduire le méli-mélo de câbles pour réduire le réchauffement de l'entrepôt tout en optimisant la circulation de l'air et la bonne gestion de son aération.

Par exemple Google a déjà opté pour cette solution en ouvrant un centre de données situé en Finlande. En 2019 il a ainsi investi 600 millions de dollars pour cet entrepôt plus écologique. Le système de refroidissement des serveurs se fait uniquement avec de l'eau de mer glaciale, une ressource entièrement naturelle.

La consommation d'énergie peut être réduite de 10 à 15 % par centre, grâce à l'autoscale et à la réduction de la mise sous tension des serveurs, lorsque le trafic est moins élevé.

Nous pouvons aussi favoriser l'utilisation de l'énergie verte qui est produite par les énergies renouvelables, que ce soit pour l'alimentation comme pour le refroidissement. Bon nombre d'entreprises souhaitent recourir à ce type d'énergie provenant des éoliennes, des centrales hydrauliques et aussi des panneaux photovoltaïques.

L'utilisation de l'intelligence artificielle serait un atout afin de déterminer avec précision les moyens nécessaires afin d'optimiser la performance énergétique tel que le taux d'humidité et la température.

Le remplacement des installations et des serveurs pourrait être faite par des solutions simplifiées et plus récentes car aujourd'hui les serveurs sont suréquipés en matériels à plus 77 % augmentant indûment leur consommation, les installations n'ont pas été créées en prenant en compte le respect de l'environnement. Dans ce même optique le stockage sur cassettes LTO réduit l'émission de CO2 de 95%. Pour donner un ordre d'idée sur une heure de fonctionnement les cassettes LTO permettent de réduire 57% des émissions de CO2 dû aux HDD en remplaçant que 60%.

80% des données archivées et 57% des données répliquées détenues par les entreprises étaient transférées sur des cassettes, les émissions de CO2 seraient réduites de 43,7% En 2030, si les entreprises transfèrent 80% de leurs données archivées et 57% des données répliquées étaient transmises sur ces disquettes nous réduirions les émissions de CO2 de 664 millions de tonnes (consommation annuel énergétique de 80 millions de foyer) ou une réduction de 43,7%.

Les solutions de réduction de la consommation énergétique ne sont pas seulement à prendre par les entreprises d'hébergement et les grandes firmes, elle doit être la priorité et l'engagement de chacun. Nous pouvons tous agir pour l'environnement et à notre échelle cela passe par la gestion et suppression de nos actions sur le net en limitant le taux d'information mis en ligne, en réduisant le nombre d'inscription réalisé sur des sites internet ou encore supprimant les historiques de navigation, les caches et les cookies stockés, en supprimant les mails qui ne sont pas importants en réalisant un tri régulier de toutes nos boîtes mail, vider la corbeille. Nous pouvons aussi nous désinscrire des newsletters pour lesquels nous ne sommes plus intéressés, leur demander de supprimer nos données. Nous pouvons fermer les pages de navigation ouvertes ainsi que fermer les applications sur son ordinateur mais aussi sur son téléphone.

La société Qarnot Computing a fait preuve de très grande ingéniosité en créant des radiateurs-numériques dédiés au particulier. Chacun d'entre nous pourrait troquer son radiateur contre ces

radiateurs qui ne sont d'autre que des ordinateurs ou serveurs réalisant le stockage, l'analyse de donnée et des calculs informatique afin d'utiliser la chaleur dégagée par les appareils pour le chauffage d'un logement.

Conclusion

L'information de masse stockées et analysées par les entreprises représente un avantage de taille face à ses concurrents. L'information et l'utilisation d'outil d'aide à la décision lui permet d'être bien renseigné sur son environnement, ses clients, ses fournisseurs tout en s'assurant de la véracité de l'information et la fiabilité de la source de l'information. Détenir de l'information pour une entreprise ne suffirait il faut d'abord qu'elle soit analysée par le biais de nouveaux outils capable d'analyser un volume d'information conséquent aujourd'hui nous parlons d'information de masse chaque outils et objet connecté relève des données pour beaucoup de donnée personnelle de toutes sortes qui sont conservé et analysé afin d'être utilisé par la suite par les entreprises afin de les aider à prendre des décisions pour leurs stratégies futures. Le volume de donnée ne cesse d'augmenter ce qui demande donc une consommation énergétique et un taux d'émission de CO2 plus important chaque jour.

En vue des problèmes environnementaux que nous rencontrons déjà actuellement nous ne pouvons-nous permettre d'augmenter cette surconsommation il est urgent de changer mœurs et les habitudes de chacun et aussi en adoptant des technologies qui permettent de réduire au maximum l'impact de la Big Data sur la planète. Ceci est une responsabilité sociale et la responsabilité de chacun en tant qu'entreprise et surtout en tant qu'individus à chacun des actions doivent être menée car les déchets électroniques s'accumulent. Il faut lutter contre des projets futurs aberrant par exemple lorsque la société Orange projette d'investir 135 millions de dollars dans un centre de donnée en Égypte ce qui ne permettra aucunement de réduire la consommation de son système de refroidissement.