Un livre blanc NetApp + APY + LEXISTEMS.

Valoriser les données d'entreprise en automatisant l'Intelligence Artificielle et le Traitement du langage Naturel

Les données sont devenues l'actif le plus précieux des organisations modernes. Dans une société globale où le numérique est désormais un acquis, les données sont à la fois l'enjeu des échanges et la base sur laquelle la valeur se crée.

Pourtant, les données sont toujours considérées comme un problème. Leur nature intangible et leur croissance exponentielle nécessitent un management adéquat – humain, technique, organisationnel… - de sorte qu'elles restent associées à des notions de coût, de complexité et de disruption.

Mais les choses changent. Pour un cas d'usage récent, LEXISTEMS, NetApp et APY se sont attaqués à une problématique typique des organisations riches de données en automatisant tout le traitement de la data par l'IA.

Acteurs clés : la solution SensibleData de LEXISTEMS et le stockage intelligent Data Fabric de NetApp, combinés dans une solution dédiée conçue et construite par l'AI Lab de APY.

Résultat : données et documents deviennent une source durable de profit, de performance et de satisfaction client -- avec à la clé de réels avantages compétitifs et la révélation d'opportunités d'affaires insoupçonnées.

La data : un défi pour l'entreprise

Les chiffres de croissance des données en entreprise sont vertigineux. Sur les 40 trillions de gigaoctets (*i.e.* zettaoctets, soit 10^21 octets) aujourd'hui stockés, 90 % ont été produits au cours des deux dernières années. C'est comme si chaque personne sur la planère produisait 1.7 mégaoctets chaque seconde. Dans le même temps, l'usage d'Internet génère à lui seul 2,5 exaoctets (*i.e.* 10^18 octets) chaque jour - fériés y compris. Sur cette base, le trésor mondial des données devrait atteindre 175 ZB en 2025.

Données utiles vs «black data» – Combien de ces données sont réellement utiles ? Selon IDC, le pourcentage de data à valeur business a grimpé à 37 % en 2020 (contre 22 % en 2012). Dans le même temps, IDC estime que seul 0,5 % de ces data sont analysées (3 % étant simplement taguées), alors que TRUE Global Intelligence estime les « dark data » (informations non-quantifiées et/ou inutilisées) à plus de 50 % en moyenne quelle que soit l'organisation – y compris la vôtre. Si, à en croire The Economist, la data est l'or noir du XXIème siècle, c'est là un énorme gâchis.

Le problème des silos – Outre les problèmes de volumétrie, la disponibilité même des données est souvent réduite du fait de structures organisationnelles en silos. Selon leur service, les collaborateurs ont des besoins en données différents, et il en va de même des clients. Chacun s'accorde généralement à dire qu'une base unifiée de données et de documents serait bénéfique à tous mais il semble qu'une force obscure empêche que de réels liens internes soient créés entre services et départements, que des synergies soient cultivées et que la valeur de l'information soit factorisée. Quand les données ne sont pas disponibles, elles ne peuvent pas être consommées.

Une infinité de formats – Du *legacy* au mobile, les données présentent une large diversité de formats et de structure, ce qui les rend d'autant plus difficiles à unifier. Plus les sources et les applications sont nombreuses - corporate, métier, e-mails, réseaux sociaux… - plus le problème est complexe, comme en témoignent les difficultés reportées par les gouvernements du monde entier pendant la récente pandémie du Covid-19. Idéalement, les données devraient être accessibles par leur contenu, indépendamment de leur contenant ou de toute autre considération technique.

Données publiques et données « open » – Il en va de même concernant les données dites « open ». Administrations et agences gouvernementales ont généralement pour mission d'ouvrir leurs données au public. Ce trésor d'informations pourrait être combiné aux données d'entreprises et transformé en *business insights* à haute valeur ajoutée. En 2020, l'interopérabilité, les cycles de vie et la langue des données ne devraient plus être des facteurs bloquants.

Bien qu'ils soient clairement documentés, ces « pain points » relatifs aux données se retrouvent dans la plupart des organisations, avec pour conséquences des pertes significatives d'opportunités et de revenus. Pour que son potentiel soit pleinement réalisé, la data doit être facilement trouvable, accessible, interopérable et réutilisable. Par les hommes comme par les machines.

SensibleData®
by LEXISTEMS :
quelques cas d'usage
ici en version mobile.

Le sens à la place des mots-clés

Dans toute organisation, les données sont de deux types : structurées (typiquement les bases de données) ou non-structurées (typiquement les fichiers PDF et Office, les e-mails et les contenus multimédia). Pour les utilisateurs, cela ne devrait pas avoir plus d'importance que le lieu ou la technologie de stockage. Les données existent donc chacun doit pouvoir les trouver immédiatement y accéder directement et travailler avec sans contrainte. Pour cela, les applications doivent être capables de comprendre parfaitement à la fois les requêtes des utilisateurs et les données ciblées. C'est là où l'ancienne technologie des mots-clés pose un problème.

Les mots-clés sont obsolètes – Historiquement, l'industrie des données s'est construite sur les mots-clés, mais les besoins actuels rendent ces derniers clairement obsolètes. Tout simplement parce que les mots-clés sont des séquences de caractères fixes, monolingues et dénuées du moindre sens, que les ordinateurs, téléphones mobiles et assistants traitent sans aucune valeur ajoutée. Du point de vue de l'utilisabilité des données, cela signifie :

- aucune variation orthographique ou imprécision
- des résultats médiocres avec les interfaces vocales
- aucun résultat sur les synonymes ou les équivalents sémantiques
- pas de questions complexes car plusieurs mots-clés se téléscopent ou s'éliminent mutuellement
- aucune adaptabilité à l'information multilingue.

Les applications fonctionnant par mots-clés sont donc limitées par nature aux traitements de données les plus basiques : elles ne font que manipuler des groupes de lettres, sans prise en compte du sens ou du contexte. Ces défauts réduisent considérablement les possibilités opérationnelles, frustrent les utilisateurs et, au final, empêchent la création de valeur.

Le sens révèle les données – Avec le sens, ces limitations disparaissent. Grâce au sens, les applications travaillent par idées et concepts, comme le font les humains, sur les données existantes comme sur les nouvelles. Elles supportent l'infinie variété de nos expressions, y compris les questions floues, imprécises, complexes ou conceptuelles. Des requêtes traitées par le sens donnent des résultats pleins de sens (c'est-à-dire plus riches et par nature contextuels), de sorte que les possibilités en matière de traitement et de connexion de données deviennent illimitées. En permettant l'exposition et la consommation d'information en langage naturel, le sens rapproche enfin les données et les utilisateurs.

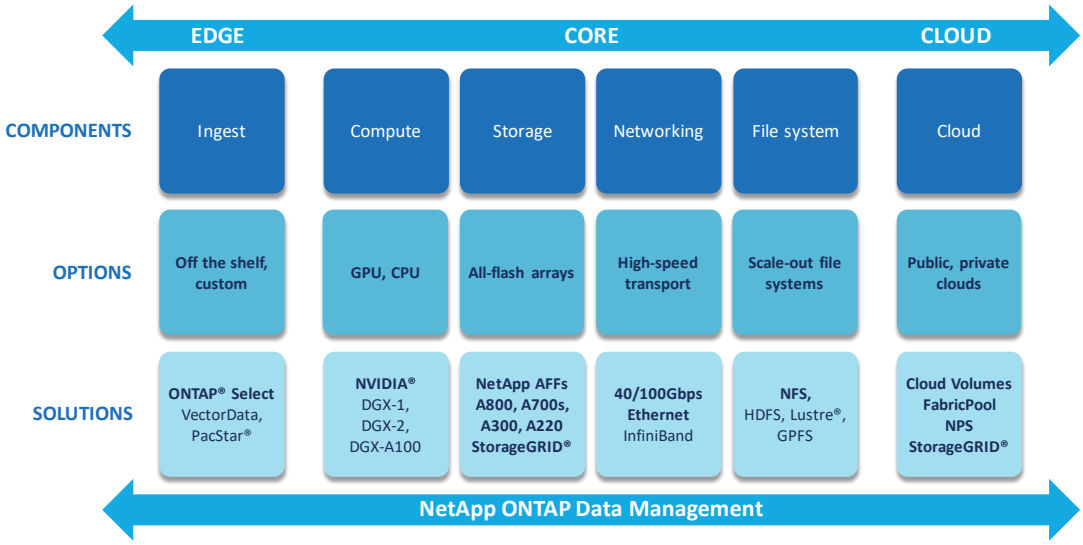
Apprentissage puissance 10 – Autre bénéfice crucial du traitement par le sens : les applications voient leur capacité d'apprentissage décuplée. Grâce au sens, les algorithmes d'apprentissage s'adaptent finement à la nature et à l'évolution des données cibles (concept et l'expression). C'est ce qui explique cette faculté propre aux applications « douées de sens » d'absorber tout de suite les vocabulaires corporate, métiers ou propres aux 'utilisateurs. C'est ce qui leur permet de rendre les données consommables automatiquement dès qu'elles sont produites.

La valeur de l'information dépend de ce que ses connsommateurs peuvent en faire. Contrairement aux mots-clés, le sens permet aux applications de comprendre les données et les utilisateurs, donc de livrer des résultats intelligents et contextuels sur toutes sortes de requêtes business. Dès qu'elles savent où sont stockés données et documents, les applications basées sur SensibleData peuvent initier un cycle continu d'apprentissage automatique en temps réel qui rend l'information consommable et monétisable comme jamais auparavant.

LEXISTEMS : les données par le sens avec SensibleData®

Pour LEXISTEMS, mots-clés et Intelligence Artificielle sont contradictoires. De ce constat est né Sensible.ai®, une technologie très innovante permettant le traitement des données et le pilotage des systèmes par le sens, en multilingue si nécessaire. Fruit de 10+ années de recherche au plus haut niveau, Sensible.ai est au coeur de plusieurs solutions d'entreprises, parmi lesquelles LEXISTEMS' SensibleData.

Contrairement aux meilleurs moteurs de recherche ou plateformes de données, SensibleData permet à chacun de rechercher, de traiter et de connecter des informations par le sens et non par mots-clés. Depuis tout type d'application et avec une totale conformité RGPD et CCPA sur les données et documents privés des organisations. En mode texte, vocal ou machine--machine, SensibleData tranforme immédiatement tout actif data en une source de profit durable et mésurable. En savoir plus...



Data Fabric by NetApp : implémentation de l'offre globale.

Stockage intelligent = données intelligentes

Appliquer l'Intelligence Artificielle aux données d'entreprise comme le fait SensibleData de LEXISTEMS nécessite un *workflow* robuste : consolidation et préparation des sources, définition et entraînement des modèles, déploiement en production et monitoring des résultats en temps réel. Avec des données partout et nulle part, sous toutes formes imaginables et dont le volume augmente en permanence, l'automatisation de l'Intelligence Artificielle, tant pour la conception des modèles par les développeurs que pour la consommation des résultats par les utilisateurs devient un réel challenge technique. D'où le besoin d'un stockage sans couture, activement intelligent, dans tous les environnements applicatifs possibles.

Le concept de « data fabric » – Unifier les données situées dans les clouds publics, privés et sur site requiert une nouvelle approche non disruptive du stockage, à la fois dans son implémentation et dans sa gestion. Cette « data fabric » (données sans couture) comprend d'une part une API agnostique de l'environnement qui exécute les même services *software-defined* quelle que soit l'infrastructure et ses fournisseurs, et d'autre part une interface utilisateur unifiée permettant la découverte immédiate, l'intégration, l'optimisation la protection et la migration des données vers / depuis n'importe quel cloud. Simple et robuste, cette architecture est en phase avec les priorités IT car elle tranforme les clouds en des extensions efficace et transparentes de tout centre de données, à volonté, en quelques clics.

Optimiser la mobilité des data – Le déploiement de modèles d'Intelligence Artificielle orientés data consiste en trois phases correspondant chacune à la localisation des données : périphérie (ingestion et consolidation des données), coeur (entraînement, *data lakes*) et cloud (archivage). Cette mobilité de la data, typique des applications d'aujourd'hui, est cause d'innombrables goulets d'étranglement. Grâce à l'approche Data Fabric de NetApp, toutes les composantes nécessaires à l'accélération puis à l'optimisation de cette mobilité s'intègrent naturellement pour offrir une performance de bout en bout incluant la préparation en local, le calcul clusterisé, l'architecture des systèmes de fichiers, la topologie réseau et le provisionning des ressources cloud, quelles qu'elles soient.

Des bénéfices immédiats – Une fois les premiers éléments de ce stockage sans couture en place, le pipeline automatique d'Intelligence Artificielle produit ses premiers résultats. A la périphérie, les échanges sécurisés de larges fichiers typiquement séquentiels bénéficient d'une très faible latence. Au niveau *core*, l'unification des datalakes (NFS, HDFS, S3, Hadoop, Splunk, Lustre, GPFS…) maximise la performance des échanges en fonction des types de data (fichiers PDF ou Office, graphes de connaissances, logs, séries temporelles, multimédia…). En fin de course, les dépôts DevOps rendent les modèles IA éligibles à l'intégration et au déploiement continus, ce qui permet soit la mise en production automatique soit l'itération cyclique jusqu'à qualification. Les coûts sont considérablement diminués sur tout le cycle de vie de l'application, de même que les délais d'IT et de Data Science.

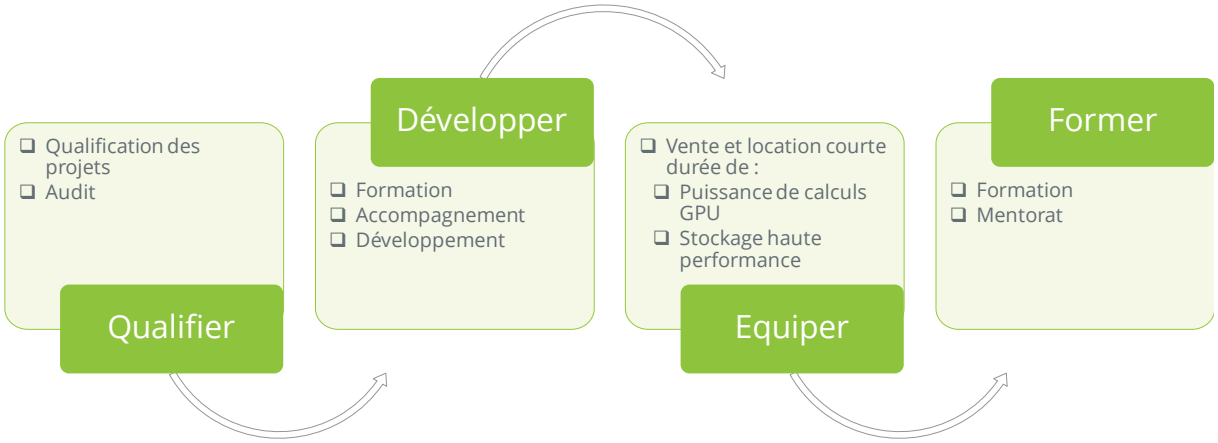
Le stockage d'aujourd'hui ne se limite plus à la persistence des fichiers, quelle qu'en soit la complexité. Dans des contextes d'applications multisources où la data est de plus en plus distribuée et hiérarchisée, le stockage intelligent Data Fabric de NetApp unifie et optimise l'accès aux données sur l'ensemble de l'environnement IT de l'entreprise. Lui seul permet d'industrialiser la production continue d'informations basées sur l'Intelligence Artificielle comme le propose SensibleData.

NetApp : Automatiser l'IA dans tout environnement de stockage grâce à Data Fabric

Alors que les fournisseurs de stockage se focalisent soit sur la charge de travail soit sur l'optimisation de la fourniture de données, NetApp unifie les deux avec Data Fabric, une architecture fondatrice qui provisionne, administre et exécute les applications de développement, de test et de production là où cela a le plus de sens à l'instant T. Grâce à une API compatible tous clouds, les utilisateurs de Data Fabric contrôlent, orchestrent, optimisent et sécurisent toutes les données à partir d'une même interface unifiée.

Cette interface intégrée est ce qui rend possible l'automatisation de l'Intelligence Artificielle. Ingestion, entraînement, production, archivage, tout est accéléré, avec une accessibilité maximale garantie - à la fois aux développeurs et aux utilisateurs - tant en local qu'à partir des data centers ou de tout cloud dans le monde entier. En savoir plus...

Ai Lab by APY : une offre à 360°, de la conception à la production.



Des implémentations spécifiques à chaque cas d'usage

Dès que votre projet de valorisation des données par l'IA commence à prendre forme, son implémentation requiert la plus grande attention. Les spécialistes sont unanimes : l'architecture et l'exécution techniques sont des facteurs clés de réussite. De là l'importance d'un partenaire réellement spécialisé en IA, capable de garantir à la fois les meilleures pratiques et de proposer les meilleures configurations.

Apprendre avec les GPUs – Les applications modernes doivent pouvoir apprendre afin de s'améliorer automatiquement et de livrer de meilleurs résultats à mesure qu'on les utilise. C'est là que l'Intelligence Artificielle entre en jeu. Reste à savoir quel type d'apprentissage. En pratique, tout dépend de la nature des données. Le *Machine Learning* donne d'excellents ré-sultats avec des données structurées – typiquement des enregistrements de bases de données ou des documents JSON stockés dans des datastores. Pour les données non-structurées – typiquement des fichiers – le *Deep Learning* est plus indiqué car les multiples couches qui composent le réseau de neurones compensent l'absence de structuration. Dans les deux cas, les GPUs NVIDIA sont les meilleurs composants de calcul.

Inférer avec des serveurs multi-noeuds – Le processus d'apprentissage génère des modèles. Ces modèles sont chargés par les applications IA pour permettre l'inférence, c'est à dire la livraison de résultats à partir de ce qui a été appris. Par rapport à l'apprentissage, l'inférence nécessite nettement moins de puissance de calcul, et peut éventuellement être déchargée pour partie sur les terminaux utilisateurs. De ce fait, les serveurs multi-noeuds à base de CPUs professionnels Intel® Xeon® constituent l'option la plus rentable pour dimensionner le système par rapport au nombre d'utilisateurs. En fonction des spécificités du cas d'usage et de la taille des modèles, les CPUs dédiés à l'inférence seront utilement secondés par des capacités mémoire de l'ordre du teraoctet, grâce notamment à la nouvelle technologie mémoire Optane® d'Intel et à des bibliothèques logicielles optimisées pour l'inférence.

Implémentation et exécution – La taille et les spécifications des serveurs hébergeant les GPUs et les CPUs – qu'ils soient « standard catalogue » ou construits sur mesure en fonction des besoins de l'application – dépendent principalement de la volumétrie des données et de la base d'utilisateurs. A cet égard, tout comme aucun projet de données n'est identique, il n'existe pas de formule universelle. En revanche, le niveau de performance ressenti est fonction de la proximité entre utilisateurs, sources de données et serveurs applicatifs - le principe étant de réduire au maximum toute latence entre questions et réponses. Un partenaire digne de ce nom saura proposer les meilleures combinaisons de clouds privés et publics afin d'assurer la livraison automatique de modèles d'inférence à l'application et par voie de conséquence les meilleures données à leurs consommateurs.

Pour ambitieux que cela semble, valoriser les données d'entreprise en automatisant l'Intelligence Artificielle et le Traitement du Langage Naturel est désormais possible. En associant le traitement des données par le sens de LEXISTEMS et le stockage intelligent de NetApp dans une infrastructure scalable, l'AI Lab d'APY a réussi à créer une application aux performances et à la rentabilité sans équivalent -- source durable de création de valeur et de satisfaction utilisateurs.

AI Lab by APY : un guichet unique pour simplifier l'adoption de l'intelligence Artificielle

Créée en 1998 et opérant sur l'Europe et l'Amérique du Nord, APY est un constructeur reconnu de solutions informatiques sur mesure pour l'industrie du Media & Entertainment et pour l'Intelligence Artificielle. Au sein d'APY, l'AI LAB aide les entreprises à concevoir les meilleures solutions pour leurs processus spécifiques et les accompagne pendant tout le cycle de vie du projet.

De l'idée à l'industrialisation, en passant par la formation des utilisateurs et la conduite du changement, les experts de l'AI Lab s'appuient sur un écosystème de partenaires privilégiés dans les domaines du Machine et du Deep Learning. Pour les clients APY, cette alliance garantit les meilleurs résultats possibles, sous la forme de solutions dédiées construites à partir d'environnements logiciels éprouvés et fonctionnant sur des équipements de calcul et de stockage certifiés. En savoir plus...