Comment construire et exploiter pleinement votre stratégie data ?

Le volume de données générées dans le monde explose. Une mine d'or pour les entreprises qui sauront les exploiter et un facteur de compétitivité critique pour les startups… à condition d'avoir su mettre en place une stratégie data efficace.

Temps de lecture : minute

---

10 mars 2022

180 zettaoctets à l'horizon 2025, un chiffre démesuré qui pourrait même être multiplié par 45 d'ici 2035. C'est tout simplement le volume de données dans le monde. Et pourtant, encore aujourd'hui, 85 % des projets data échouent, en partie du fait d'une mauvaise gestion ou d'une gouvernance inadaptée. Face à ce constat, déployer une bonne gouvernance des données n'est pas une option, c'est une question de survie. Alors, quels sont les freins potentiels ? Comment faire face à une montée en puissance des données au sein de son activité ? Quel accompagnement choisir ? Quelles sont les clés d'une stratégie adaptée à son business ?

Concentrez-vous sur vos enjeux data spécifiques

C'est la première question à vous poser pour établir votre stratégie data. Difficile en effet de faire les bons choix stratégiques quand on ne connaît pas ses besoins prioritaires. L'agence de voyages en ligne Combigo, par exemple, combine différents moyens de transport (avion, bus, train) afin de dénicher les trajets les plus économiques et à l'empreinte écologique la plus faible. Un positionnement original qui implique d'agréger un

maximum de données, de les contextualiser et de les rafraîchir constamment pour disposer des informations à jour.

" *La stratégie comme la gouvernance des données sont spécifiques à chaque entreprise. C'est pourquoi la première étape consiste à identifier précisément les besoins data à appréhender en priorité et qui vous permettront de vous démarquer,* explique Vincent El Khatib, cofondateur et CTO de Combigo. *Dans notre cas, notre principale problématique repose sur l'ingestion des données : comment formaliser la donnée brute et la restituer pour nos algorithmes. Grâce à eux, nous examinons en moyenne près de 6 000 documents par seconde pour détecter la meilleure combinaison possible.* "

Utilisez le bon outil pour passer à l'échelle

Pour faire face au déluge de données, Combigo a choisi d'utiliser MongoDB, une base de données orientée document qui permet de modéliser ses fichiers et autres schémas sous la forme souhaitée. À partir de là, le choix des serveurs d'OVHcloud semblait évident pour Vincent el Khatib : " *La proximité réseau entre nos instances de calcul et la base de données nous permet d'obtenir les meilleures performances possibles. Et nous avons la garantie que tous les backups sont sauvegardés dans un autre datacenter pour des raisons de sécurité.* "

Autre point important pour le CTO : les développeurs de Combigo n'ont plus à réaliser les montées en version de la base de données MongoDB eux-mêmes, puisque OVHcloud assure la migration d'une version à l'autre. Un gain de temps qui permet aux équipes de procéder par itérations rapides pour croître plus rapidement. Et comme la startup doit stocker entre 400 et 500 millions de documents sur ses serveurs Cloud, ce n'est pas une mince affaire.

À lire aussi
Entraîner son IA dans le Cloud ? Oui, s'il est français

Identifiez vos freins potentiels

À l’image des itérations qui vous permettent de changer rapidement de stratégie, il est essentiel d’identifier les freins qui risquent de vous ralentir. Là aussi, les difficultés dépendent essentiellement du positionnement de votre entreprise et de sa taille. Une étape d’autant plus importante que dans une petite structure comme une startup, qui compte peu de ressources, ces obstacles peuvent vite s’avérer difficiles à contourner. Dans le cas de l’entreprise Combigo, les freins rencontrés se positionnent essentiellement à deux niveaux.

" *Le premier concerne notre capacité à faire s’accorder la donnée entre le moment où elle est disponible et celui où l’utilisateur effectue sa requête,* détaille le cofondateur. *Le prix des billets aériens par exemple évolue constamment, ce qui complexifie la manière dont nous devons structurer la donnée. Nos algorithmes font alors une re-tarification pour estimer le prix du billet et rendre la donnée viable le plus longtemps possible. Notre deuxième difficulté majeure repose sur la très grande quantité de*

données à analyser, ce qui ralentit le processus d'extraction des informations pertinentes qu'elles contiennent. Il faut penser à mieux segmenter et cibler la donnée dès le départ pour aller droit à l'essentiel. "

Misez sur les bons indicateurs

Une fois vos premières itérations lancées, pensez à mettre en place des indicateurs précis et fiables qui vont vous permettre d'analyser vos tests. Cette étape de mesure des performances vous donne également l'opportunité de déterminer les métriques manquantes, et de compléter si besoin à chaque test pour aller de plus en plus loin et obtenir une meilleure compréhension de votre business.

" *Essayez d'être le plus MVP* [ndlr : produit minimum viable] *possible. Inutile de coder tout le produit dès le départ. Commencez par une petite partie et si ça marche, alors allez au bout des choses. C'est de cette manière que vous serez efficace. L'essentiel, lors de chaque test, est de progresser sur le binôme temps passé – apprentissages reçus pour vous permettre d'avancer en continu.* " Une clé de succès indispensable pour Vincent El Khatib.

Déléguez les tâches à faible valeur ou trop pointues

Difficile de tout faire en interne lorsque l'on est une startup. D'autant plus que certains profils tech se font rares sur le marché. La bonne nouvelle, c'est que l'écosystème de la French Tech ne manque pas de partenaires pour vous accompagner, à l'image d'OVHcloud et de son Startup Program. L'intérêt est double pour les startups : d'une part, elles répondent à leurs besoins data du moment, et d'autre part, elles accélèrent leur développement, en France comme à l'international.

" *L'accompagnement de leurs architectes techniques nous a permis de*

tester de nombreuses hypothèses très rapidement et de scaler plus sereinement, admet le CTO de Combigo. *En déléguant à OVHcloud toute la partie Infrastructure, nous avons ainsi pu nous concentrer en interne sur nos développements et nos bases de données, tout en profitant de la puissance d'automatisation de leur cloud. Résultat, en quelques clics à peine, nous avons pu doubler notre volumétrie de données entre décembre et janvier.* "

Quels profils externaliser et lesquels privilégier en interne ? La question est difficile à trancher mais certaines expertises peuvent plus facilement être confiées à des tiers. Parmi lesquelles la cybersécurité qui exige des compétences de pointe, mais aussi la capacité à interroger une base de données (*requêtage en langage informatique)* et l'organisation des données, tâches particulièrement chronophage. À l'inverse, les profils experts en traitement de la donnée – *data engineer* et *data scientist* en particulier – sont plus simples à internaliser.

Anticipez votre croissance… mais pas trop

Dernière étape : préparez l'avenir et pensez *scalability*, c'est-à-dire comment intégrer la croissance de vos volumes de données au bon moment pour accompagner le développement de votre startup. Un point important, mais attention car anticiper la croissance trop tôt peut aussi devenir un frein. Passer à l'échelle reste en effet un enjeu critique dans la vie d'une startup. Par conséquent, cette étape peut s'avérer assez longue et complexe à gérer. En l'envisageant trop tôt, vous risquez non seulement d'affecter votre activité en cours, et de ralentir votre développement, mais aussi de générer des surcoûts inutiles liés au surdimensionnement de vos serveurs cloud. L'accompagnement d'un partenaire peut là aussi vous permettre d'agir au bon moment et de bien préparer le sujet.

" *Commencez par valider une hypothèse et réfléchissez au fur et à*

mesure à la scalabilité. Allez au maximum de vos possibilités et une fois arrivé à un goulot d'étranglement, alors là, passez à l'échelle. Le cloud vous permettra de scaler en quelques minutes autant de fois que nécessaire ", confie Vincent El Khatib.

Un dernier conseil ? Collectez un maximum de données, quitte à en avoir trop et à ne pas toutes les utiliser. Stockez-les et voyez plus tard comment les exploiter au mieux. Une précaution qui peut vous permettre de prendre une longueur d'avance sur vos concurrents le moment venu.

Maddyness, partenaire média d'OVHcloud.

Article écrit par Maddyness, avec OVHcloud