5 conseils pour monter la meilleure équipe de data scientists

Les profils de data scientists sont parmi les plus recherchés par les startups. Mais un data scientist ne ressemble pas à un autre. Renaud Allioux, co-fondateur de la startup Preligens, donne ses astuces pour repérer et convaincre les meilleurs talents.

Temps de lecture : minute

---

24 janvier 2022

En tant que CTO d'une entreprise d'intelligence artificielle, j'ai passé du temps à structurer l'équipe tech et réfléchir aux profils à embaucher lors des différentes phases de croissance. D'abord, au lancement de l'entreprise, à la définition du minimum viable product, puis lors de la phase de passage à l'échelle et du lancement des produits en production chez les clients. Si vous êtes CTO, vous allez avoir besoin d'un certain nombre de data scientists mais tous n'auront pas les mêmes missions. Voici leurs différents profils et ce que j'ai compris en formant mon équipe IA à Preligens qui comprend aujourd'hui plus de 70 data scientists pour 200 employés.

Conseil n°1 : un data scientist doit savoir coder !

Ce n'est pas négociable. C'est mon mot d'ordre et même la chose la plus importante que j'ai apprise durant mes 5 années de CTO. Tous doivent être en mesure d'intégrer leurs modèles dans un cadre afin que celui-ci soit utile à l'application et au cas d'usage. Oubliez les *Jupyter notebooks*.

Si vous voulez dépasser l'expérimentation et proposer de réels produits d’IA, vos data scientists doivent savoir coder et proposer des produits concrets.

Conseil n°2 : pour la première embauche, choisissez un Full Stack Data Scientist

Généralement utilisé pour les développeurs, le terme *full stack* convient également pour qualifier un data scientist qui possède un panel de compétences plus large que le plus grand mathématicien théoricien résolvant les problèmes les plus difficiles avec une craie et un tableau noir. Un véritable couteau suisse qui doit présenter :

- de solides compétences informatiques,
- un bon niveau d’ensemble en mathématiques,
- du pragmatisme et de l’audace,
- des bases de DevOps (Docker, CI/CD) et visualisation,
- idéalement 3 à 5 ans d’expérience.

Conseil n°3 : lors de la constitution de votre première équipe, recrutez des data scientists spécialisés

Lorsque l'équipe passe d'un à 4-5 data scientists, c'est généralement le moment où votre premier produit sera suffisamment mature pour être mis sur le marché. Dans cette équipe, vous avez besoin d'un ensemble complémentaire de compétences et de profils. Il s'agit toujours de data scientists, mais :

- certains plus orientés vers la science,
- certains plus spécialisés sur l’approche itérative de R&D,
- et d’autres davantage tournés vers le codage.

Encore une fois, tous, absolument TOUS doivent être capables de coder correctement et d'intégrer leurs modèles. C’est cette combinaison de compétences qui vous permettra d'atteindre vos premiers jalons en termes de développement de produits.

Conseil n°4 : tous les data scientists doivent savoir coder... mais ils ne veulent pas s’occuper de l’annotation des données

Le meilleur moyen de voir un data scientist quitter votre équipe est de lui faire labelliser lui-même ses données. Vous devez vous doter d'une équipe dédiée, d'outils et de processus pour le Data Labeling afin qu'ils n'aient pas à le faire. Un data scientist heureux veut utiliser les données qui sont déjà étiquetées ; sans cela, il sera frustré et partira vers une autre entreprise !

Conseil n°5 : structurez les équipes au moment de passer à l'échelle

Chez Prelignes, nous avons structuré notre équipe de data scientists en 3 parties différentes :

1. L’équipe de production IA

Ils sont chargés de l’exploitation de notre stack IA pour produire des applications pour les clients. Ils sont pragmatiques et itératifs, font de la R&D continue. Ils ont principalement des formations *full stack* ou *computer vision*. Ils font un peu de tout et sont multidisciplinaires. Tournés vers le client, ils sont également bons en gestion de projet et en communication. C’est un mélange de jeunes diplômés et de leads expérimentés qui forment ensemble notre “première ligne”.

2. L’équipe ingénierie IA

Ils sont chargés du développement, de l'amélioration et de la maintenance de notre pile d'IA. Ils sont dévoués et méticuleux, ils construisent le moteur de l'usine. Ils ont un très haut niveau de codage, ainsi que de solides connaissances en DS. Ils fournissent des piles sur lesquelles les autres équipes peuvent travailler mais les forment également pour qu'elles soient autonomes dans l'ajout de fonctionnalités. Ce sont généralement des personnes plus expérimentées que la production et les meilleures sur le plan technique.

3. Les chercheurs IA

Ils sont chargés de résoudre des problèmes à haut risque et à fort potentiel. Bien qu'ils soient très scientifiques, ils doivent également comprendre le code pour s'assurer que les innovations sont mises en production. Ils sont généralement titulaires d'un doctorat en mathématiques, en biophysique ou en astrophysique. Ils proposent une exploration scientifique pour résoudre les problèmes de production mais produisent également des articles de conférence ou de journal. Tous ont *a minima* 3 à 5 ans d'expérience.

La combinaison de ces trois équipes n'est pas d’un laboratoire mais une véritable usine, capable de livrer des produits qui sont déployés en production chez les clients et utiles pour des cas d’usage dans leur quotidien.

Diplômé de Centrale Supélec et titulaire d'un doctorat en astrophysique, Renaud Allioux a cofondé Preligens en 2016 avec Arnaud Guérin pour développer des solutions d'intelligence artificielle pour des applications critiques. Il dirige une équipe technique de plus de 120 collaborateurs. Il est passionné par la mise en application de l'IA, le passage à l'échelle des

équipes et des processus pour atteindre l'excellence technologique.

Article écrit par Renaud Allioux