

Data Scientist : longue pénurie annoncée pour un nouveau métier

1 – Le point de vue français...

Pourquoi un nouveau besoin et un nouveau profil ?

Pierre Morizet-Mahoudeaux – Face à l'explosion des volumes de données, les investissements en datacenters pour le traitement et le stockage et en datawarehouse tentent de donner du sens à toute cette information.

De nouvelles ressources et technologies comme le Big Data (mais aussi la virtualisation, le cloud, etc.) permettent d'en tirer profit. Et les appels d'offres se multiplient concernant ces sujets.

D'où le besoin croissant en recrutement de data scientists : du spécialiste des données jusqu'au stratège d'entreprise comprenant les métiers et l'analyse des données. Les premiers correspondent à un profil d'ingénieur de type master (Bac+5), tandis que les seconds proviennent plutôt d'écoles de commerce avec une spécialisation sur les données. Les entreprises devront donc choisir le plus spécialisé avec cette double casquette, ou un tandem efficace.

En effet les analyses de données sur lesquelles se basent les technologies de type Big Data (mais pas uniquement) nécessitent cette double casquette données/métiers.

Autre besoin : le statisticien. En effet, les modèles statistiques prédéfinis engendrent souvent trop de possibilités d'interprétation. Pour résoudre ce problème, des spécialistes doivent pouvoir affiner les modèles afin d'obtenir du sur-mesure indispensable.

Les formations informatiques préparent-elles réellement les étudiants à ce type de fonction proche des métiers des entreprises ?

Une bonne formation informatique confère aux étudiants une base pérenne sur laquelle ils peuvent toujours s'appuyer pour s'adapter aux situations : méthodologies, algorithmes, une ou plusieurs plates-formes qui illustrent ces principes... En outre, les masters ouvrent sur des spécialisations.

Enseigner les nouvelles technologies pour les nouvelles technologies présente un intérêt très limité, et ne rend pas forcément opérationnel.

Ainsi, un de nos étudiants en stage chez EDF pour des programmes de tests liés aux capteurs intelligents n'a rencontré aucun problème pour utiliser Hadoop, alors qu'il n'y avait pas été initié.

Comment se traduit l'implication métier ou la spécialisation sur le terrain de la formation ?

Tout d'abord, notre établissement a noué de forts partenariats avec des grandes entreprises

informatiques (éditeurs et SSII), dont les spécialistes dispensent des cours : SAP, Teradata, Sas...

Pour favoriser l'approche métier et scientifique, 30 % du temps de nos formations d'ingénieur sont consacrées à autre chose que la science technologique : banque, finance, sociologie, gestion, innovation...

Une préorientation spécialisée est choisie dans un domaine dès le début des études, avec la possibilité de modifier cette orientation par la suite. En fin de cycle, l'ingénieur informatique est donc opérationnel dans le domaine choisi. D'autant plus que les stages sont également choisis dans cette optique.

Un étudiant peut terminer sa formation avec une spécialisation sur les données, la stratégie d'entreprise et le datawarehouse, mais aussi sur les statistiques et les mathématiques, ou encore sur l'infrastructure réseau.

2 – ... en accord avec le point de vue américain.

Quelle est la situation aux États-Unis sur cet aspect ?

Stephen Brobst – On assiste à une forte demande en data scientists.

Les universités et écoles américaines conçoivent des programmes en intégrant plus de sciences et de mathématiques. Une démarche initiée encore insuffisante sur les aspects statistiques, visant surtout à former des personnes opérationnelles sur le marché du travail.

Cependant, de nombreux aprioris ont la vie dure, comme la perception du geek, ou le fait que les maths sont plutôt pour les garçons. Pourtant, les quelques femmes présentes dans ces métiers sont souvent excellentes.

Autant de circonstances qui laissent présager que le fossé entre le besoin et l'offre va encore se creuser.

Comment identifiez-vous et qualifiez-vous ce type de profil ? Informatique ou pas ?

Pour répondre aux attentes analytiques et autres Big Data, les entreprises ont besoin de profils connaissant à la fois les données et les métiers.

Cependant, autant un spécialiste des données peut se spécialiser assez simplement à des métiers, autant le contraire est plus qu'improbable.

Autre distinction importante, le data scientist n'est pas un informaticien (computer scientist). Pour lui l'informatique peut éventuellement lui fournir un outil, et va déployer de façon informatique ce qu'il a conçu.

Par exemple, les langages statistiques comme R sont très différents de langages de programmation comme Java ou C#.

D'ailleurs, la collaboration entre ces deux fonctions est généralement limitée à la mise en production. Ainsi, il revient à l'équipe informatique de développer les applications ou déployer un

cluster Hadoop, par exemple.

Le data scientist joue un rôle d'intermédiaire entre les métiers et l'informatique.

Crédit photo : © pro motion pic – Fotolia.com

Voir aussi

[Quiz Silicon.fr – Le vocabulaire du cloud](#)