

Recruter un data scientist ? Bienvenue au Far-West

La bulle n'est pas encore prête à se dégonfler. En fin de semaine dernière, se tenait à Paris la **seconde édition de DataJob**, manifestation consacrée aux métiers de la Data. Une manifestation qui cherche encore son rythme de croisière, mais qui a vu affluer nombre d'étudiants en recherche de stage ou d'emploi. Nul doute que les **salaires mirobolants** accordés aux data scientists, ces spécialistes de l'interprétation des Big Data, ainsi que les progressions de carrière rapides (on devient senior dès 2 ou 3 années d'expérience) y sont pour beaucoup. De facto, la 'science de la donnée' (ou Data Science) reste un domaine mal balisé pour les recruteurs. « *Parmi les personnes se prétendant Data Scientist, on trouve des écarts de 1 à 100 en termes de compétences entre un néophyte et un expert* », illustre **Gilles Babinet**, digital champion français auprès de la Commission européenne. Et de parler d'une « *phase un peu délirante* » en ce qui concerne les salaires. Une vision que confirme **Jérémy Harroch**, fondateur et Pdg de la société de conseil Quantmetry, à l'origine de DataJob. Dans nos colonnes, [ce dernier parlait il y a quelques jours d'une « bulle »](#) entourant les data scientists.

(En complément notre infographie : [PHP, .Net, Java, iOS, Big Data : le salaire des développeurs en 2014](#))

Des tâtonnements somme toute logiques quand on se penche sur le profil franchement atypiques de ces professionnels. « *Un métier à l'intersection des maths, de l'IT, du Machine Learning et des statistiques. Le tout avec une forte orientation métier, analyse Réda Gomery*, directeur associé chez Deloitte. *C'est un mouton à 5 pattes qui reste difficile à trouver d'autant que la formation française a tendance à privilégier les cursus spécialisés* ». Même si des **formations dédiées** mettent sur le marché du travail leurs premières promotions d'experts, ces profils restent rares. Et ces formations encore jeunes. Or les besoins sont déjà là. Le directeur associé chez Deloitte explique ainsi voir émerger « *de plus en plus de projets – dans le marketing, les ventes, la lutte contre la fraude ou dans la cyber-sécurité – directement liés à la valorisation des données ou à l'exploitation des techniques analytiques* ».

« Surmonter les défis de la banque française »

Présent sur DataJob, **la Société Générale recrute** ainsi 2 Ddata scientists senior et de deux junior pour renforcer son équipes de datamining (une trentaine de personnes, experts techniques, spécialistes des risques ou du marketing). « *Nous cherchons des profils qui savent aller identifier et apprivoiser de nouveaux outils pour les intégrer à l'existant. Un peu dans l'esprit de l'Open Source. Être un utilisateur des logiciels phares du domaine ne suffit plus* », tranche **Joseph Trojman**, directeur de la stratégie et des études à la Société Générale (banque de détail). Des profils qui collent avec les nouveaux besoins de l'équipe que dirige Joseph Trojman : « *Les clients attendent de l'immédiateté et de la personnalisation. La question du multicanal devient donc centrale puisque cette dernière implique de bien comprendre le parcours de chaque client* ». Pour l'équipe de dataminers de la SG, c'est un changement profond. Plutôt que de travailler sur des moyennes et des variances, il faut désormais se concentrer sur de nouvelles modélisations. Joseph Trojman assure que, au sein de sa banque, cette transformation a démarré voici deux ans, et qu'elle aboutit aujourd'hui à des **résultats qualifiés de**

« **prometteurs** ». Pour le directeur de la stratégie et des études, « *les analyses prédictives servent à surmonter les challenges que doit affronter le modèle relationnel bancaire français* », un modèle de plus en plus bousculé par des clients enclins à faire jouer la concurrence.

Insuffisant pour Gilles Babinet, qui explique que les grandes entreprises françaises font majoritairement fausse route : « *au sein du CAC 40, on trouve souvent des équipes Data Science isolées alors que le Big Data appelle une logique transversale* ». Selon une étude de l'éditeur SAS Institute, plus de 50 % des data scientists aux Etats-Unis sont stressés souvent en raison d'un **positionnement défaillant dans leur organisation**. « *Plusieurs facteurs expliquent ce mal-être : des problèmes de communication dans l'entreprise, la définition assez floue du rôle de data scientist, une fonction associée à d'importantes responsabilités mais sans autorité et le manque de compétences sur les technologies employées* », détaille **Mouloud Dey**, directeur des solutions métiers et des marchés émergents au sein de l'éditeur d'outils statistiques.

« La bataille de la data a commencé »

Pour Gilles Babinet, cette absence de maturité se reflète dans les **attributions encore balbutiantes des premiers CDO** (Chief Data Officer, responsables des données). « *Le Chief Data Officer a un rôle diplomatique au moins aussi important que la maîtrise du digital. Sa fonction n'a rien à voir avec celle du Data Scientist où domine le besoin d'expertise verticale. Et c'est là où le bât blesse aujourd'hui. Les vrais Chief Data Officer n'émergeront que dans quelques années* », tranche le digital champion. Aux Etats-Unis, l'évolution semble enclenchée : l'étude de SAS Institute note la montée en puissance graduelle de profils moins techniques, plus proches de la stratégie ou de l'organisation, au sein des équipes de Data Science.

CDO de la France (ou plus exactement [administrateur général des données de l'Etat](#), selon la terminologie officielle), **Henri Verdier** (en photo en haut de l'article) se veut plus optimiste. Selon lui, les révolutions technologiques (la micro, Internet, le mobile...) sont relativement déconnectées les unes des autres. Nulle raison que le Big Data n'échappe à la règle. « *On commence la bataille de la data ; les compteurs sont remis à zéro* », assure-t-il. Façon de dire que l'Hexagone a ses chances, d'autant que l'enseignement en mathématiques y est réputé. Et d'expliquer que **l'administration française aurait déjà acquis un certain leadership** mondial en la matière, grâce à l'écosystème existant en France autour de la donnée. Et le directeur de la mission Etalab (qui pilote la politique Open Data de l'Etat) de mettre en avant un certain nombre de projets lancés ou réalisés au cours des 12 derniers mois : **Open Fisca** (simulation de droit fiscal et social), **base nationale d'adresses** (géolocalisation de tous les bâtiments de France), **Pole Emploi Store** (exposant les ressources sous forme d'API), **stratégie d'Etat plate-forme** ou encore le **registre national des taxis** (indiquant leur localisation en temps réel). « *Que le législateur régule ce marché, secoué par l'arrivée d'un acteur comme Uber, via un projet portant sur les données sous-jacentes à cette activité est une idée fondatrice que, j'espère, on retrouvera dans d'autres domaines* », dit Henri Verdier.

De la ministre du Logement à la médaille Fields

Et ce dernier d'insister sur les basiques, les fondamentaux de la Data Science par lesquels il faudrait démarrer toute réflexion. « *La Data Science ira jusqu'à la médaille Fields (le plus prestigieuse récompense*

mondiale en mathématiques), mais démarre par des choses simples. Avant tout, il faut **s'attaquer au problème de la non-circulation de la donnée**, dit l'administrateur général des données de l'Etat. *Que la ministre du Logement ne dispose pas en temps réel des prix de l'immobilier est problématique. Il faut commencer par amener les données brutes aux gens qui en ont besoin* ». Un prérequis qu'illustre Joseph Trojman (Société Générale) : « la multiplication des contacts digitaux avec les clients génère de nouvelles données que les banques doivent apprendre à exploiter. Car, contrairement aux acteurs de la grande distribution, elles n'avaient pas cette habitude. » **Objectif** : consolider ces nouvelles données avec les informations plus classiques (souscriptions de produits, rendez-vous en agences...) sur une échelle des temps mettant en évidence le « *comportement relationnel du client* ».

Henri Verdier y ajoute une dimension : la philosophie d'utilisation de ces données : « *faut-il mettre la Data Science au service de la concentration des pouvoirs ou au contraire au service du plus grand nombre via une stratégie d'empowerment ?* », lance l'administrateur général des données de l'Etat. La Data aurait donc **une portée politique...** interrogeant le modèle hiérarchique même des entreprises.

A lire aussi :

[Marketing et DSI s'écharpent autour du Big Data](#)

[Big Data : « Une bulle autour des Data Scientists »](#)