

# Randstad : le Big Data pour gommer le paradoxe du marché de l'emploi

Le Big Data au service de l'emploi ? L'idée n'est pas vraiment nouvelle, mais Randstad (3 600 personnes en France, 2,7 milliards d'euros de chiffre d'affaires en 2014) a le mérite de pousser le concept assez loin, grâce à l'exploitation de technologies avancées de Machine Learning. « *La solution que nous avons développée vise à répondre au paradoxe de l'emploi en France, martèle d'emblée le président du spécialiste de l'intérim dans l'Hexagone, François Béharel (en photo). Le pays compte 4 à 5 millions de chômeurs et des acteurs de l'emploi comme nous ont du mal à servir les commandes des entreprises* ». Avec des pénuries sur certains profils (comme les soudeurs) et/ou sur certains bassins d'emploi particuliers. Selon le dirigeant, Randstad ne parvient ainsi pas à satisfaire 1 commande sur 2 ou sur 2 ½, un ratio qui varie en fonction de la saison et des profils.

C'est sur la base de ce constat que la société s'est lancée, il y a trois ans, dans un projet de Big Data. Avec la volonté de dépasser le simple outil de mise en correspondance, via analyse sémantique, des offres et des CV disponibles. Au total, la solution développée par Capgemini gère quelques 1 000 qualifications, mais surtout 11 000 compétences différentes. « *En cas de pénurie sur une qualification, on peut ainsi zoomer sur les compétences nécessaires pour élargir la recherche et trouver des passerelles entre les métiers* », illustre François Béharel. Une piste permettant, par exemple, de proposer une formation à un métallier afin de le muer en soudeur.

## **Hadoop + Spark**

Basé sur l'exploitation des 3 millions de CV que possède le spécialiste de l'intérim et des 12 millions d'annonces d'emploi publiées par an que la société récupère auprès de partenaires (Pôle Emploi mais aussi les jobboards), l'outil repose sur des technologies d'analyse sémantique et sur un arbre de décision permettant d'identifier les compétences acquises (dans les CV) ou attendues (dans les offres), ainsi que leur poids (correspondant, par exemple, au nombre d'années correspondantes dans le CV). C'est ici que sont intervenues les technologies de Machine Learning, mais également l'expertise des équipes de Randstad sur le sujet. « *C'était, avec l'acquisition de données, le gros morceau sur ce projet* », dit Valérie Perhirin, qui dirige les activités Big Data et analytics de Capgemini France. Démarrés en 2013, les développements reposent sur Hadoop (un cluster de 100 To sur base de distribution Cloudera), un développement Spark – aujourd'hui le framework star pour le système de gestion de données Open Source – et sur une infrastructure Oracle.

## **Efficacité commerciale et nouveaux marchés**

Centré au départ sur l'algorithme de rapprochement des compétences et des profils – le volet qui doit avoir un impact direct sur l'efficacité commerciale des consultants Randstad –, le projet s'est peu à peu enrichi d'autres services, tournés vers les collectivités territoriales ou la direction des entreprises. De facto, l'outil, présenté ce matin par le DSIO (directeur de l'organisation et des systèmes d'information) du groupe, Christophe Montagnon, propose des scénarios d'usage très

divers. Certains tournés vers les candidats (orientation, entrée sur le marché du travail, anticipation des reconversions), d'autres vers les recruteurs (disponibilité de compétences sur leur bassin d'emploi, élargissement de la recherche en cas de pénurie). Mais l'outil offre aussi une vision stratégique aux entreprises (disponibilité de compétences dans le temps sur un bassin d'emploi afin de cibler une implantation, concurrents sur un bassin d'emploi en terme d'attractivité sur un métier donné) mais aussi aux collectivités ou à l'Etat (attractivité des territoires, bonne utilisation de l'argent public). Christophe Montagnon a ainsi présenté une carte permettant de mesurer, département par département et qualification par qualification, la tension entre l'offre et la demande. En somme, Randstad voit dans sa solution Big Data un outil permettant de se différencier de ses concurrents, tant en améliorant l'efficacité de ses commerciaux qu'en le positionnant sur de nouvelles problématiques.

Aujourd'hui testé au sein d'un projet pilote auprès d'une population experte, l'outil doit prochainement être amené aux 2 000 à 2 500 consultants de Randstad, déjà équipés de tablettes. Christophe Montagnon explique que ces utilisateurs, qui ne bénéficieront que d'une partie des visualisations que permet l'outil, seront équipés en octobre. Randstad entend par la suite enrichir son outil de modèles prédictifs et envisage aussi d'y agréger la base de CV détenue par Pole Emploi. Une version publique, accessible à tous, est également à l'étude. Cette dernière devrait proposer une vision statistique, François Béharel ayant clairement indiqué sa volonté de « *ne pas brader cet outil* » en donnant accès à tous au savoir-faire de son groupe, « *à commencer par la passerelle entre métiers et compétences* ».

*Mise à jour le 16/09 à 12h40*

Selon Oracle, la taille du cluster Hadoop atteint 100 To et non 10 To comme nous l'avions écrit (cette donnée a été corrigée dans l'article). La solution repose sur une appliance Big Data de l'éditeur américain intégrant la distribution Cloudera. Le cluster est lié aux bases de données de Randstad fonctionnant sur Oracle Exadata.

#### **A lire aussi :**

[Comment Monster affine le recrutement grâce au Big Data](#)

[Comment Criteo transforme Hadoop en moteur de sa rentabilité](#)

[Hadoop sur le déclin ? Cloudera, Hortonworks et MapR répondent](#)