

## [Tribune] Big data : en finir avec les idées reçues

Simon Ouellet, co-fondateur de MiniBig

Qui dit big data dit big entreprise? Non, rétorque Simon Ouellet. Le co-fondateur de MiniBig revient sur les idées reçues les plus fréquentes sur le big data.

Les technologies de traitement de données connaissent un développement inédit et font régulièrement la une de la presse grand public. Mais le traitement journalistique qui en est fait ne reflète pas tout à fait la réalité et la diversité de ce domaine d'activité en constante évolution. Pour les professionnels du secteur, il est temps de prendre la parole et de s'attaquer à quelques idées reçues particulièrement répandues.

" Le big data ne sert qu'au marketing "

Les applications les plus connues des technologies de big data concernent en effet le traitement des données personnelles par les géants du web qui tirent la majorité de leurs revenus de la publicité. Du fait de la focalisation des médias sur cette partie émergée de l'iceberg, le public a acquis une certaine méfiance vis à vis de ces techniques devenues synonymes de manipulation et d'espionnage généralisé.

Mais les applications du big data sont infiniment variées et ne se limitent donc pas au marketing. De la sociologie à la médecine, de nombreux champs de recherche ont ainsi été révolutionnés. La principale avancée permise par ces techniques est la sortie de la double ignorance : " on ne sait pas que l'on ne sait pas ". Le traitement de données à grande échelle permet de détecter des corrélations sans en chercher une en particulier. Une étape décisive de la recherche, qui consiste à chercher les questions pertinentes à poser avant vérification empirique, a donc été automatisée et optimisée. De nombreux programmes de recherche ont été lancés sur cette base, sans faire la une des médias pour autant.

" Le big data ne sert que les grosses entreprises "

Autre conséquence de la surmédiation de leur usage par les géants du web, le big data est perçu comme un outil réservé aux grands groupes. Pourtant, ces technologies se démocratisent extrêmement rapidement grâce notamment à des prestataires spécialisés qui mettent à disposition de clients de toute taille des ressources de calcul et des programmes d'analyse en SaaS.

Chez les professionnels, cette idée reçue trouve son origine dans le difficile changement de paradigme que représente le big data. Les DSI sont habituées aux architectures centralisées et monolithiques, dont le coût est particulièrement élevé, mais les acteurs les plus rodés au big data ont depuis longtemps adopté un modèle basé sur un grand nombre de matériels de taille moyenne reliés en un vaste réseau distribué. La force du big data ne repose pas sur la disponibilité et la fiabilité mais sur la multiplication. C'est ce décalage entre un nouveau modèle et des DSI encore trop timides qui pousse de nombreuses entreprises à faire monter inutilement les budgets de leurs infrastructures.

Pour en finir avec ces idées reçues, il est nécessaire de communiquer mais aussi d'agir et de changer concrètement les pratiques. L'image du big data est entachée de nombreuses atteintes à l'éthique par des entreprises traitant les données personnelles sans considération pour leurs propriétaires. Il est primordial d'adopter la transparence comme ligne de conduite et de prévenir les utilisateurs des services en ligne de tous les usages qui sont faits des informations les concernant. Ce n'est pas le traitement des données qui dérange les internautes mais bien l'opacité qui l'entoure. N'attendons donc pas que le législateur s'empare de la question et impose des règlements - qui se révéleront rapidement inapplicables ou inadaptés - pour améliorer réellement l'image de nos métiers !

