

<http://www.20minutes.fr/monde/1038034-20121108-election-americaine-big-data-arme-secrete-obama>

Election américaine: «Big data», l'arme secrète d'Obama

HIGH-TECH La collecte et le traitement automatisé de gigantesques quantités de données a aidé le président à remporter un second mandat. Et les experts à prédire avec une précision chirurgicale les résultats...

Philippe Berry

Publié le 08/11/12 à 00h00 — Mis à jour le 17/11/16 à 16h37



Capture d'écran du site BarackObama.com. — DR
[De notre correspondant aux Etats-Unis](#)

En 2008, l'équipe d'Obama avait pris une longueur d'avance sur le terrain des réseaux sociaux, utilisés pour lever des fonds et coordonner la logistique de terrain. Cette fois, la crise et le chômage avaient remplacé l'espoir et le changement. Et pourtant, malgré un léger effritement du vote populaire, le président a réussi à presque faire aussi bien. Son avantage sur Romney? Un torrent de données sur les électeurs, les supporters et les Etats clés. Time a bénéficié d'un accès exclusif au QG de Chicago et [lève un coin du voile](#) sur les coulisses des opérations «Narwhal» et «Dreamcatcher».

100 analystes dans «la cave»

«Big data», pour des quantités d'information tellement vastes qu'elles ne peuvent pas être traitées par l'humain ni par les bases de données habituelles. A la place, il faut une puissance de stockage et de calcul massive, ainsi que des informaticiens et des statisticiens pour faire chanter les données.

Après la campagne de 2008, le camp Obama décide de quintupler le nombre d'analystes «maison», à presque 100, [selon Mother Jones](#). Le patron, Jim Messina, très proche du président de Google, Eric Schmidt, recrute deux nerds. Rayid Ghani, un ancien d'Accenture, chercheur spécialisé en «data mining et machine learning», utilisait notamment ses talents pour maximiser l'efficacité des promotions des supermarchés. Spécialiste de l'analyse «en temps réel», Harper Reed, lui, s'est fait les dents sur le réseau des transports publics de Chicago. L'équipe de choc est réunie dans une pièce sans fenêtre baptisée «la cave».

Un chiffre derrière chaque décision

Etape 1: amasser les données. Les sources principales: les registres électoraux publics, les affiliations au parti démocrate ou républicain, les profils des supporters sur MyBarackObama et ObamaForAmerica et, grâce à Facebook Connect (utilisé pour se connecter sans créer un profil), les informations que Mark Zuckerberg accepte de partager avec des partis tiers.

Etape 2: interconnecter les données. En 2008, chaque pôle de la campagne (collecte de fonds, porte à porte, spots publicitaires) disposait de ses informations, parfois sous forme de registres papiers. Cette fois, une seule gigantesque base a été constituée, accessible à tous les niveaux, notamment sur smartphone et tablette, pour une action coordonnée.

Etape 3: utiliser les données. Obama a bombardé ses supporters d'emails ciblés les appelant à donner quelques dollars. Exemple type: un fan écrit sur le site officiel que les questions sociales sont particulièrement importantes pour lui, il reçoit un email sur les dérapages des républicains sur le viol et l'avortement. L'efficacité de chaque courrier était mesurée, notamment l'impact du nom de l'expéditeur (Michelle Obama, Joe Biden etc.). Publicité, budget alloué à chaque Etat... Chaque décision était basée sur un chiffre. L'élection était encore simulée 66.000 fois chaque soir avec les derniers chiffres. Le matin, la stratégie et les ressources étaient ajustées en conséquence.

Des données parasitées

En face, Romney n'est pas resté à l'âge de pierre. Le stratège Karl Rove a été l'un des pionniers d'une stratégie basée sur les nombres. Mais selon Mother Jones, au moins jusqu'aux primaires, la campagne républicaine a surtout outsourcé cette mission au lieu d'en faire le cœur de sa machine. Surtout, les modèles républicains de la réalité sur le terrain se sont révélés incorrects, avec de nombreux Etats qu'Obama a largement remportés alors que Mitt Romney y avait activement fait campagne.

«Le challenge principal, c'est de trouver des données neutres. Certains conseillers politiques parasitent la sélection des données par leur intuition», explique à 20 Minutes le neuro-scientifique Sam Wang, qui a pronostiqué l'élection avec une efficacité de 100% au sein du [Princeton Election Consortium](#). La morale: «Les ordinateurs ne font du bon boulot que si on leur fournit de bonnes informations.»

Big data ou le triomphe des prévisions

Le grand vainqueur de cette élection, c'est Nate Silver, le statisticien du blog [FiveThirtyEight](#) (comme les 538 votes du collège électoral), hébergé par le New York Times depuis 2010. Si la Floride est déclarée en faveur d'Obama, Silver aura prédit avec justesse les résultats dans les 50 Etats. L'informaticien a fait ses classes dans la prédiction des performances des joueurs de baseball avant de passer à la politique. Il a construit un modèle en s'appuyant sur les chiffres de milliers de sondages qu'il passe à une moulinette secrète. Sa victoire est surtout un triomphe sur les «pundits», ces experts en sciences politiques qui sont souvent plus proche d'un historien que d'un informaticien. Le républicain Michael Barone, qui misait sur un facteur non rationnel –la plus faible mobilisation des minorités– a reconnu sa défaite, expliquant qu'il pouvait «manger son chapeau». Comme Mitt Romney.

MOTS-CLÉS :

- [Etats-Unis Présidentielle Barack Obama Élection présidentielle américaine](#)
- [maison blanche mitt romney](#)