

Big Data et Marketing Politique:

Presse et Medias

Le présent rapport parle uniquement de quelques élections clés aux Etats-Unis, en France et le Brexit est parfois mentionné, et par conséquent uniquement les problèmes soulevés lors de ces élections sont mentionnés, cette restriction est nécessaire, le sujet étant si vaste qu'il est impossible de le présenter et résumer dans son ensemble.

Le Big Data suscite beaucoup d'engouement dans le domaine scientifique, et d'aucuns s'y plaisent à y voir l'avenir de la science informatique et mathématique.

Que se cache-t-il dans sous ce nom ? Le Big Data consiste en la collecte, le traitement et l'analyse/interprétations de données massives (plusieurs milliards de milliards d'octets). Le terme est apparu pour la première fois en 1997 et les applications du big data n'ont cessé de se diversifier et de se développer dès lors.

L'avènement d'Internet et de quelques entités s'étant rendus inévitables (Facebook, Google, les fournisseurs d'accès Internet,...) a servi de terreau au développement du Big Data, permettant une collecte de données sans précédent dans notre histoire. Les données obtenus sur Internet ont prit une telle ampleur que cela est devenu un *business*, de nombreux data brokers sont alors apparus sur le marchés, vendant au plus offrant leurs immenses bases de données (Axciom, Experian, InfoGroup mais aussi Facebook, Google eux-mêmes,...). Ces bases de données sont ensuite analysées par les acheteurs qui en déduisent des stratégies marketing pour convaincre leurs clients/electorat/cibles.

Si l'emploi du Big Data pour le ciblage publicitaire est connu depuis des années, une nouvelle tendance s'est développée en 2008 et 2012 et transforme depuis le paysage : l'emploi du big data en politique.

Difficile de dire à quand remonte la première utilisation du big data en politique mais ce qui est certain c'est la date de sa popularisation, de son « sacre » international : l'élection en 2008 et la réélection en 2012 de Barack Obama.

C'est ainsi que fut projeté sur le devant de la scène l'emploi massif des données à des fins de marketing politique.

Dès lors le développement fut exponentiel. Mitt Romney, Ted Cruz, Hillary Clinton, Barack Obama, Bernie Sanders, Nicolas Sarkozy, François Hollande, Bruno Le Maire, Nathalie Kosciusko-Morizet, Jacques Cheminade, Emmanuel Macron, François Fillon, Jean-Luc Mélenchon,... Tous ont en commun deux choses : la politique et l'emploi du big data dans leurs campagnes.

Ce Duo (Big Data et Marketing Politique) fut donc remarqué par la presse en 2012 majoritairement.

2008 n'était que les prémices, le vague murmure d'un changement au sein même de notre conception de stratégie politique mais 2012 fut la consécration, la preuve aux yeux du monde que le big data n'était pas qu'un simple gadget technologique, de la poudre de perlimpinpin si j'ose dire.

Avec un intérêt accru dans la presse internationale à partir de 2012 le big data s'est peu à peu inscrit dans une controverse à mi-chemin entre celle dont elle faisait déjà l'objet (principalement sur la protection des données personnelles, de la vie privée) et celle bien connue de la politique qui s'incarne parfaitement en la citation de James Freeman Clarke : « La différence entre l'homme politique et l'homme d'Etat est la suivante : le premier pense à la prochaine élection, le second à la prochaine génération » . On parle alors dans le domaine d' « homme politique sismographe ».

En effet si je voudrais simplifier énormément en catégorisant des camps je dirais que l'on peut distinguer le camp des « pour » , ce sont principalement les utilisateurs et les créateurs, les architectes du big data, les politiques, et le camp des « contre » ce sont en particulier les victimes du big data, les pourvoyeurs de leurs propres données personnelles, les défenseurs de la démocratie.

Mais la réalité est bien plus complexe.

Tout d'abord ,si vous le voulez bien, divisons ce problème en deux sous-problèmes : le premier lié à l'emploi en général du big data et son lien avec la vie privée, la collecte de données, le second en lien avec les conséquences de telles méthodes pour la vie politique, pour les valeurs du fonctionnement démocratique.

Dans le premier cas on peut dire que jusqu'à l'arrivée du Big Data dans l'arène politique il y avait stabilisation. En effet une sorte de *statu quo* s'était instauré, d'une part grâce à la régulation imposée par les institutions tels la CNIL en France, (exemple rapport de la CNIL sur le Safe Harbor). D'autre part à cause de la menace pesant sur l'image des entreprises auprès de l'opinion publique (qui s'imposent alors un cadre restrictif dans leur collecte et analyse) mais également aussi par l'attrait du Big Data permettant une source de revenus pour les entreprises et assurant alors la gratuité de certains services quasi-indispensables comme Google, Facebook,... et permettant de fournir une offre sur-mesure à chaque personne (services,offres,contenus, publicités personnalisés/adaptés).

L'arrivée du big data dans le domaine politique est venue chambouler ce statu quo, comme si l'on avait réveillé les vieilles blessures, un nouveau *casus belli* mêlé de confusion.

Casus belli car par la suite, les medias, la presse se sont mis à commenter allègrement chaque mouvement , chaque nouveauté utilisant les mêmes faits au service d'argument différents.

Une chose est certaine, il y a transformation du paysage politique et c'est l'origine de notre problème comme le résume si bien la Stampa dans un article du 23 septembre 2012, un des premiers sur le sujet :

« Big Data è la fine della società che eleggeva presidenti come Roosevelt o Nixon, un messaggio per tutti, diffuso dai mass media. Big Data è la comunicazione «personal» in una società dove contano gli individui, ciascuno diverso dagli altri, Libertà, Fraternità, Diversità. Big Data sa quel che tantissimi leader e analisti stentano a capire in Italia, che ci sono elettori «di destra» su tasse e bilancio e «di sinistra» su matrimoni e adozioni gay,

che ci sono elettrici «progressiste» sulla spesa da aumentare nella scuola, ma «conservatrici» sulla spesa da augmenter nella difesa. »

comme le dit le journaliste Gianni Rotta, le Big Data signifie la fin des projets pour tous, des messages qui s'adressent à l'ensemble des citoyens en tentant de les rassembler derrière un idéal, une idée commune. Au contraire le Big Data accepte et joue sur notre diversité en tant qu'individus : on peut être « conservateurs » sur le budget de la defense et « progressiste » sur celui de l'éducation. Le Big Data permet de formuler des messages individuels, sur-mesure. Chacun est différent des autres. Liberté, Fraternité, Diversité.

C'est justement cette transformation, ce passage du message pour tous au message à chacun, qui est le coeur du notre problème.

Dans une interview pour GQ Magazine, Vincent Pons (fondateur de LMP l'entreprise d'analyse et recueil de données utilisée par Emmanuel Macron principalement avec son algorithme « 50 +1 ») déclare :

« Un fossé se creuse entre les politiques et une partie de la population qui se sent à l'écart. Le porte-à-porte est justement un moyen pour les politiques d'aller directement écouter les électeurs »

LMP a en effet recueilli plus de 25 000 témoignages grâce au porte à porte mais également des statiques INSEE et bien d'autres données encore.

Cet argument de permettre de toucher plus facilement les électeurs, d'être à l'écoute de leurs préoccupations est repris dans e-marketing (un journal en ligne spécialisé dans les technologies du marketing numérique mais également dans tout le domaine du démarchage, des stratégie de vente) :

"C'est une façon de maximiser la prise en compte du plus grand nombre de préoccupations partagées par les citoyens »

« Le porte-à-porte est plus rentable que le reste selon le livre de référence du marketing politique, Get out the vote !, d'Alan Gerber et de Donald Green: vous pouvez faire changer d'avis un électeur sur 14, contre un sur 38 au téléphone et un sur 100 000 par mail! Quant aux réseaux sociaux, ils permettent le plus souvent de mobiliser son électorat plutôt que de l'étendre »

Les Echos dans un article du 19 avril 2017 s'interroge sur l'efficacité de cette nouvelle tendance en politique :

« Les sociétés qui vendent ce type de logiciels avancent que cela peut faire gagner 1 à 2 points à un candidat »

« Vu l'état de volatilité très forte de l'électorat, le potentiel de l'efficacité de l'hyperciblage permis par le Big data n'a jamais été aussi élevé en France », estime Benoît Thieulin, fondateur de l'agence de communication digitale Netscouade devenu directeur de l'innovation d'Open.

« Ce qui est frappant dans cette élection, c'est qu'une partie des candidats communique beaucoup là-dessus pour faire innovant, moderne. Alors qu'auparavant, la question de

l'utilisation des données personnelles était jugée sensible par tous et toutes et en restait au stade des coulisses », souligne Anaïs Théviot, chercheuse en sciences politiques et auteure d'une thèse sur le militantisme partisan en ligne.

« D'une certaine manière, les partis ont toujours travaillé sur la data, à partir, par exemple des adresses postales puis ensuite de simples fichiers Excel. Mais aujourd'hui, c'est beaucoup moins artisanal et plus professionnel », poursuit-elle.

Dans Libération on peut lire :

« La question aujourd'hui, c'est celle de la mobilisation. Obama, lui, a fait du profilage. En France, ni le droit ni la culture n'y prédisposent, ça créerait d'énormes problèmes de perception par l'opinion. »

En général les « pro-Data » justifient leurs pratiques et leur engouement par les bénéfices que peut apporter une analyse des données : être plus proche du peuple, de ses préoccupations, rendre plus efficace une campagne. De plus la législation imposée par la CNIL et les autres institutions en France font qu'il devient difficile de critiquer la collecte de données étant donné que celle-ci s'effectue dans un cadre restreint et clair. En effet dans le cadre de LMP par exemple ou encore l'application Knockin' utilisée par Nicolas Sarkozy les seules données utilisées sont des données publiques (tels la base de l'INSEE) et des données obtenues par du porte à porte ou des sondages. Difficile en effet de critiquer une telle approche. L'eau se trouble beaucoup plus lorsque l'on s'intéresse au géant américain NationBuilder, réprimandé en 2013 par la CNIL pour récupération abusive de données personnelles. Cette entreprise est utilisée aux USA mais également en France par Les Républicains. De plus NationBuilder fait appel à des Data Brokers, elle achète des données privées, obtenues par Facebook, Google, des boîtes mails,... ce type de méthode est largement controversé et critiqué. Outre même le cadre de la politique, cela concerne le Big Data.

Libération, qui est en tête de proue dans la question de la protection de la vie privée déclare dans un article du 23 mars 2017 :

« Entre la vie privée des citoyens et les intérêts des opérateurs de télécommunications, le Sénat américain a manifestement fait son choix. Il a en effet voté, ce mercredi, un projet de résolution visant à abroger des règles adoptées en octobre dernier par la Commission fédérale des communications (FCC), durant la fin de mandat de Barack Obama. Ces nouvelles règles imposaient aux fournisseurs d'accès à Internet (FAI) d'obtenir l'autorisation de leurs clients avant de collecter et monnayer leurs données personnelles – notamment leur géolocalisation ou leur historique de navigation – à des fins publicitaires. »

C'est ici que les « contre » entre en jeu, outre le danger pour la protection de nos données (comparaison des Data Brokers à un Big Brother 3.0) , nombreux sont ceux qui s'inquiètent de l'impact de cette « dictature » de la Data sur la Démocratie.

Un article du New York Times intitulé 'Why the NPG VAN model is hurting Democracy' dénonce le monopole de NPG VAN sur les données du parti Démocrate Américain. Selon l'auteur (qui s'avère être le directeur de la division politics and advocacy d'un autre géant des data , NationBuilder) le fait que tout un parti (aussi immense que celui des Démocrates) ne fasse appel qu'à un seul pourvoyeur de données est une erreur. Cela offre un moyen de pression sur les candidats qui se voient refuser un accès à la base de données si ils ne suivent pas à la lettre les directives du parti.

The Guardian indique également que l'ICO (équivalent de la CNIL en Angleterre) mène une enquête sur le rôle de Cambridge Analytica dans le Brexit et l'accuse d'avoir porter atteinte à la vie privée :

« The ICO spokeswoman confirmed that it had approached Cambridge Analytica over its apparent use of data following the story in the *Observer*. "We have concerns about Cambridge Analytica's reported use of personal data and we are in contact with the organisation," she said. »

Plus loin encore, on est en droit de s'interroger sur l'efficacité du Big Data sur l'évolution des mentalités , le ciblage électoral, les messages personnalisés ne permettent pas une discussion, un débat, on convainc un électeur avec des idées auquel il adhère déjà, et non à coup d'arguments défendant une nouvelle ou différente idée. Cela pousse , dans un sens positif, les politiques à faire de nombreux compromis dans leurs programmes mais également , dans un sens négatif, cela pousse ces derniers à changer leurs positions en fonction des réactions avis de l'électorat. On cherche à se rallier à la cause de l'électorat et non à tenter de rallier l'électorat à sa cause.

Enfin beaucoup, notamment des journaux comme Le Monde, Télérama ou le New York Times, s'interrogent sur l'efficacité du Big Data au sein même d'une campagne avec comme argument principal la défaite d'Hillary Clinton, dont la campagne est historiquement celle qui a le plus massivement utilisé le Big Data. Mais également en invoquant l'échec de tous les sondages ayant utilisé le Big Data pour prédire les élections , personne n'a vu venir la défaite de Clinton, ni la victoire du Brexit, ni l'absence de François Fillon au second tour.

Le Monde critique alors les échecs de Filteris qui avait prédit une victoire de François Fillon au premier tour, ainsi que les échecs Clinton dans un article : La victoire de Donald Trump ou les trois échecs du Big Data. Télérama dénonce l'échec des Data Scientist dans la campagne américaine.

La défaite de Clinton serait due à son usage du big data, se reposant trop sur les données pour la victoire, elle n'a pas réussi à mobiliser l'électorat. Tandis que Trump lui s'est plus concentré sur la communication.

Ce n'est pas la première fois que l'on oppose Clinton et Trump, Big Data et Réseaux Sociaux. Le problème est que Donald Trump a lui aussi fait massivement appel au Big Data via Cambridge Analytica une start-up qui a travaillé pour Ted Cruz auparavant et pour les pro-Brexit. Seulement ce dernier s'est fait plus discret dans son emploi du big data et s'en est servi pour connaître les points de son programme sur lesquels insister sur quel segments de la population via du porte à porte et l'envoi de mail ou appels téléphoniques.

Ainsi ce raccourci souvent observé, sur le fait que Trump n'utilise pas le Data et que c'est la clé de sa victoire, montre une chose : il semble que le recueil de données via le démarchage , le porte à porte dans le but d'envoyer des messages personnels aux électeurs soit assez bien vu dans l'ensemble tandis que l'achat de données via des compagnies privées pour ensuite modifier son programme pour convaincre un maximum d'électeurs est mal perçu.

on peut d'ores et déjà imaginer une stabilisation du problème dans ce sens.

Toute la difficulté réside en la partie cachée, ou se résume à la question : qui croire ? De nombreuses entreprises comme Orange ou Facebook ou Google essayent de montrer une certaine transparence dans les données qu'elles recueillent. La plupart des sites indiquent également si ils utilisent des cookies et l'on peut les refuser ou les accepter. Mais là où rien n'est dit c'est sur l'emploi de ces données, si il paraît acceptable et de bonne guerre que l'entreprise utilisent les données pour elles-même il en est tout autre si elle vend à n'importe qui des données. C'est aux USA que le problème est le plus important car il n'y a quasiment aucune régulation et la proximité des politiques avec le Big Data en font qu'il est difficile de voir s'élever des contres-pouvoirs. En France il y a la CNIL, en Angleterre il y a l'ICO. Mais le développement de la technologie est tel qu'il est difficile pour la législation de suivre. Il est de même difficile pour la presse et l'opinion publique de savoir de quoi il en retourne en temps réel.

Sur regards.fr on s'interroge dans un article sur les moyens de contourner cette consommation de données en partant du principe que la transformation du paysage économique est déjà présente et qu'il ne faut non pas lutter contre directement mais simplement l'éviter :

« Facebook a bien l'intention de continuer de s'approprier toutes les informations que nous fournissons sur son site.

Comment pourrait-il en aller autrement puisqu'il s'agit là du cœur du modèle économique de tous les Google, Amazon et autres Twitter de la Silicon Valley : offrir, souvent gratuitement, un service (de partage de contenus, de microblogging, de moteur de recherche...) et collecter en échange les traces des internautes pour les revendre à des partenaires commerciaux. »

l'article propose notamment d'utiliser la désactivation des cookies, l'emploi de navigateur de type Tor,...

Je dirais en conclusion qu'il est difficile de trouver de véritables positions dans ce débat, la complexité technologique et scientifique du big data couplée avec la complexité du débat sur la vie privée et la politique font qu'il est difficile de savoir qui a raison et qui a tort. Les mentalités évoluent et changent au fur et à mesure des faits. La richesse des possibilités avec le big data fait qu'il est difficile d'y être opposé sur toute la ligne (il y a une différence entre le porte à porte et la vente de vos discussions facebook). De même étant une science en plein développement il est difficile de se prononcer encore sur l'efficacité du Big Data et en particulier en terme de marketing politique (là encore certains se limitent à son emploi pour « mesurer la température » , d'autres pour contacter les électeurs et certains carrément pour changer de ligne politique selon l'interlocuteur). Sur la vie privée également difficile de se positionner car les données sont offertes avec notre consentement, si les méthodes sont discutables éthiquement la seule critique sérieuse que l'on peut faire est que l'on nous propose pas d'alternative (on est obligé de se plier aux règles car il est difficile de se passer de Google et Facebook qui sont quasiment les seuls sur le marchés) et qu'il y a un manque de transparence. C'est à chacun de se forger une opinion en pesant le pour et le contre du Big Data en politique mais pour cela il est évidemment nécessaire d'avoir des articles de qualités et plus d'informations.

