

[Compte-rendu] Theviot, Anaïs (2019). Big Data Électoral Dit-moi qui tu es, je te dirai pour qui voter. Éd. Le bord de l'eau, 197 p.¹

Cet ouvrage paru en 2019 expose les conclusions d'un total de 77 entretiens semi-directifs menés avec des professionnels de la communication numérique des équipes de campagne du PS, UMP, LR et LREM (entre 2012 et 2017), ainsi que 20 entretiens à des directeurs d'organisations de big data électoral en France, à savoir : *LiegyMullerPons*, *Apallian*, *Emakina*, *Fédéravox*, *Netscouade*, *DigitalBok*. L'objectif de l'ouvrage est de pénétrer dans les coulisses du big data électoral afin d'interroger les promesses des algorithmes, notamment, celle de remporter une élection. Theviot montre les changements observés à l'issue des campagnes politiques, où le numérique devient central mais non excluant, c'est-à-dire qu'en plus des sources des campagnes « classiques » (des données sociodémographiques, des données issues de sondages d'opinion et des données retraçant les histoires des votes), s'ajoutent les traces numériques laissées en ligne. Ces dernières semblent être les plus intéressantes à interroger à la lumière des scandales récents comme celui de Cambridge Analytica. Ce compte-rendu expose les principales idées de l'ouvrage tout en respectant la même organisation de l'auteure.

Le premier chapitre retrace de façon diachronique l'usage du numérique dans les campagnes électorales françaises à travers de nombreux exemples. Au début des années 2000 il s'agissait de transposer les tracts vers les sites pour élargir l'audience, « la communication sur Internet n'est pas la priorité des candidats » (p. 17). Pour les élections régionales de 2004, les premiers blogs politiques apparaissent. L'auteure identifie à ce moment l'émergence d'une *rhétorique de la proximité* décrite ainsi, « le candidat s'exprime alors à la première personne

¹ Camila Pérez Lagos
CIM-Communication, Information et Médias
Université Sorbonne Nouvelle-Paris 3
F-75005
Courriel : perez.lagos.camila@gmail.com/ camila-perez-lagos@sorbonne-nouvelle.fr

du singulier pour présenter son programme et les atouts de son parcours. Cela permet de rompre avec la distance classique entre gouvernants et gouvernés » (p. 22). En 2006, au sein des partis, des pôles dédiés au TIC se sont formés. Il apparaît par exemple « l'adhésion à 20 euros du PS » qui est « pensée comme un moyen d'améliorer la visibilité du parti et de coller davantage aux "attentes" de "l'électorat de gauche", mais celle-ci entre en confrontation avec les intérêts et les équilibres internes » (p. 33), à savoir le questionnement sur la véritable adhésion de ces partisans. Plus tard, dans le cadre des élections présidentielles de 2007, il s'agit de « se saisir du numérique pour militer. Le parti [UMP] sollicite ainsi régulièrement "ses supporters" pour que leurs idées politiques soient visibles sur le réseau » (p. 19). Il existe une « volonté d'engager une dynamique participative via le numérique » (p. 34) comme par exemple « désir d'avenir » de S. Royal, qui prétend co-construire le programme de la candidate. Néanmoins, l'auteure souligne que ces expériences confirment « la difficulté pour les partis politiques de se saisir des potentialités du web 2.0 (davantage centré sur les pratiques collaboratives et l'interactivité), car ils ne sauraient se détacher de leur mode d'organisation traditionnelle (volonté de contrôle et formation hiérarchisée) » (p. 38). Indépendamment de cela, internet devient un impératif pour les partis.

L'auteure expose ensuite une liste de cas qui illustrent une mauvaise campagne sur le net, comme par exemple des « lipdub » ou les « Créateurs des possibles » de l'UMP (entre 2007 et 2012), qui auraient fini par ridiculiser le parti, ou encore lorsque « l'achat de bases de données commerciales pour diffuser des messages partisans s'est accompagné de fortes critiques des internautes qui se sentent "spammés" par l'envoi massif de mails non désirés » (p. 57). Puis, à partir de 2012, Theviot identifie l'émergence d'une « rhétorique du changement et de la nouveauté » (pp. 70-71), où les partisans se rassemblent pour participer à la discussion en ligne aux moments des débats télévisés (voir le phénomène du « double écran » p. 64). L'objectif est double : attendre que le journal télévisé reprenne la parole en

ligne ce qui fait augmenter l'audience (du tweet), tout en orientant son cadrage médiatique. A partir de 2017, même si les fausses informations existent depuis toujours, ces dernières montent en puissance, « l'environnement numérique semble favoriser l'accélération et l'amplification des informations erronées, des idées complotistes, de la propagande et de la pseudoscience, ainsi que la constitution de groupes, de communautés et de réseaux voués (du moins en partie) à la diffusion de ces contenus » (p. 71). En France, l'auteure évoque les exemples d'A. Juppé comme la caricature « Ali Juppé » et la rumeur sur le comptes *offshore* de E. Macron, entre autres. En dehors de la France, les polémiques autour des élections de D. Trump et du *Brexit*, ont fait modifier à Facebook son algorithme pour « privilégier les contenus partagés par les gens "de confiance" dans le fil d'actualité des usagers » (p. 72) et s'est mis en relation avec des agences de *fact-checking* « dans le but de diminuer les contenus suspects » (p.73).

Le deuxième chapitre interroge la fabrique de données et de logiciels à des fins politiques, tout en partant du constat que la récolte des données et le ciblage est une méthode de campagne qui a toujours existé pour mesurer les options et réduire l'incertitude électorale, ce sont les outils qui changent. Inspirés des techniques du marketing, le recours aux data permet de segmenter l'électorat et de cibler les messages. Dans ce cadre, la campagne d'Obama est comprise comme le « point de départ des *data analytics* » (p. 82) ce qui a entraîné une hausse de dépenses dans les données (collecte et traitement). Néanmoins, à la différence de la campagne d'Obama, en France « l'usage des data en politique reste très dépendant des budgets » (pp. 87-88). Egalement, le fichage par ethnie ou religion est difficilement répliquable en France du fait de la protection des données. De plus, les modes d'organisation sont différents : « en France, le travail des professionnels politiques des data doit se confronter aux méfiances des cadres partisans, aux façons de faire traditionnelles et aux jeux de concurrence avec les professionnels de la communication permanents du parti » (p. 93).

La complexité de la constitution des bases de données est démontrée par divers cas comme le non partage des données des primaires de 2012 du PS ou l'essai de l'UMP en 2011 pour interdire la récolte d'emails. Lors de ce chapitre l'auteure s'intéresse, également, à l'usage des logiciels électoraux en France. Ces derniers incluent un premier type de logiciels comme *LiegeyMullerPons* et *FéféraVox*, qui permettent de croiser des données ouvertes des résultats électoraux, provenant principalement de l'INSEE ainsi que des renseignements volontairement laissés par les adhérents. Un deuxième type de logiciels comme *NationBuilder* ou *DigitaleBox* permet « d'orchestrer la campagne électorale dans tous ces aspects » (p.112) à savoir : site web, mobilisation en ligne ou hors ligne, bases de données avec degrés d'engagement, réseaux sociaux, emails, SMS, etc.

Le troisième chapitre aborde comment les campagnes électorales ont intégré le big data aux autres formes de campagnes traditionnelles comme le tractage, le porte-à-porte ou les meetings, ceci sans perdre de vue que « les big data ne font pas toute une campagne, mais participent à sa rationalisation » (p. 121). Par exemple, l'équipe d'Hollande, s'inspirant de la campagne d'Obama 2008, a cherché à cibler en amont « les zones où la proportion d'abstentionnistes de gauche est la plus forte, pour rendre encore plus efficace leur action » (p. 124), la suite consistait à faire une campagne efficace qui ne dépasse pas les 5 minutes. Or « contrairement à la campagne porte-à-porte de B. Obama, celle de F. Hollande n'a pas réussi à fédérer de sympathisants de gauche, sans lien direct avec des adhérents. En effet, seuls les sympathisants proches de réseaux militants se sont joints aux adhérents pour frapper aux portes » (p. 129). D'ailleurs, un marché autour du Big Data électoral se constitue : d'ici à 2020 les dépenses liées au big data en France « atteindraient près de 9 milliards » (p. 135) et l'enjeu « est de créer ou consolider de l'ordre de 137 000 emplois » (p. 135). De ce fait, Theviot identifie un changement concernant la qualification professionnelle, qui réside dans le fait qu'« il ne s'agit donc plus seulement de penser le web comme un outil de

communication, mais aussi comme un outil stratégique de mobilisation et d'organisation du terrain hors ligne » (pp.138-139). Conséquemment, les « travailleurs de la donnée » (p.142) cherchent à se différencier des métiers contigus, comme ceux de sondeur ou de communicant numérique. Theviot montre l'exemple de l'usage de *Nation Builder* aux primaires de la droite et du centre en 2016. L'usage de data est mis en avant par les partis en suivant la même logique de modernité ; c'est le cas du passage d'A. Juppé de *WordPress* à *NationBuilder* en 2015. Ceci a permis de créer un site web, gérer les médias sociaux, construire une base de données (comprenant celles des réseaux sociaux non autorisées par la CNIL), avoir un *backoffice* du site, gérer les SMS, la *mailing list*, les messages vocaux et automatiser les tweets. La création du parti En Marche a également été marquée par une « dimension participative » (p. 168) et en rupture avec la tradition. Néanmoins, « quand on interroge les députés LREM sur leur campagne, ils n'évoquent pas un usage des big data intense [...]. Ils préfèrent insister sur le militantisme traditionnel effectué, le contact et la proximité voulus sur le terrain avec leurs électeurs qui auraient fait la différence » (p. 175).

Lors de la conclusion, Theviot fait allusion à l'intitulé de son ouvrage tout en distinguant l'univers des possibles entre les Etats-Unis par le biais de campagnes plus « agressives et intrusives » (p. 177) et la France où, « pour le moment, on devrait plutôt dire : "dis-moi qui tu es (en laissant de multiples traces en ligne) et je t'enverrai un message ciblé pour te convaincre de voter pour mon candidat" ». (p. 177). Ceci est suivi d'une brève réflexion sur la question des données personnelles, les conseils de la CNIL pour l'application du RGPD, le scandale Cambridge Analytica et la méfiance des citoyens envers les GAFAM.

Enfin, l'ouvrage présente un travail de récolte de données et d'analyse soigneux, ce qui en fait une référence pour les chercheurs et chercheuses intéressés par la communication politique en ligne.