Anaïs THEVIOT, *Big Data électoral. Dis-moi qui tu es, je te dirai pour qui voter*

Lormont, Éd. Le Bord de l'eau, coll. Territoires du politique, 2019,
200 pages

Camila Pérez Lagos

Édition électronique

URL : <https://journals.openedition.org/questionsdecommunication/23102>

DOI : 10.4000/questionsdecommunication.23102

ISSN : 2259-8901

Éditeur

Presses universitaires de Lorraine

Édition imprimée

Date de publication : 15 novembre 2020

Pagination : 464-466

ISBN : 978-2-8143-0586-1

ISSN : 1633-5961

Référence électronique

Camila Pérez Lagos, « Anaïs THEVIOT, *Big Data électoral. Dis-moi qui tu es, je te dirai pour qui voter* », *Questions de communication* [En ligne], 37 | 2020, mis en ligne le 15 novembre 2020, consulté le 14 juin 2022. URL : <http://journals.openedition.org/questionsdecommunication/23102> ; DOI : <https://doi.org/10.4000/questionsdecommunication.23102>

---

internet fiable, omniprésente et haut-débit. Or, dans de nombreux pays, celle-ci est inexistante, ou lorsqu'elle existe, très coûteuse » (p. 151). Les avantages numériques du *cloud* sont aussi abordés : « Une réglementation des entreprises utilisatrices du *cloud* s'impose de toute urgence. Il devrait être interdit aux avocats de stocker des données clients sur le *cloud* » (p. 164). Quant à la propagande des géants du numérique, « les Gafam seraient les dignes successeurs de Marlboro dans la production d'une ignorance répandue à propos d'un défi majeur pour nos vies » (p. 170). Ailleurs, il est question de notre identité numérique : « Non seulement le contenu publié par les autres, mais également le nombre et le type de personnes qu'un individu fréquente dans son réseau changent l'identité d'un sujet en ligne » (p. 183) ; « la *privacy* assure la protection des points de vue et donc la possibilité pour un sujet de changer, de devenir quelqu'un d'autre – que cela se produise en vertu de choix autonomes ou bien d'une hybridation complexe et médiée » (p. 193). S'agissant des libertés individuelles, on peut lire que « le totalitarisme et les régimes émergents de surveillance numérique ont en commun de contraindre le sujet politique à travers des épistémologies et des technologies de prédiction, et en prétendant à l'autorité » (p. 197). Et sur l'éthique, « la vérité est un fait humain, c'est-à-dire un fait qui a à voir avec l'épistémologie en tant que connaissance humaine de l'ontologie, la vérité a toujours un aspect moral. Ce qui est épistémologiquement erroné est aussi moralement injuste. Réciproquement, celui qui a connaissance d'une vérité et se comporte, par intérêt personnel, de façon non conforme à ce qu'il sait commet une erreur morale » (p. 246). Et sur la spécificité du code numérique, il est souligné que « le code informatique n'est pas seulement un langage non naturel, c'est aussi quelque chose qui est toujours en mutation. C'est un code toujours en attente de mise à jour, et il en va de même pour les formes de confiance qui l'habitent » (p. 250).

Par ailleurs, l'ouvrage présente des mécaniques d'écriture discursive complémentaires puisque, si la majorité des auteurs choisissent de composer leur contribution comme une démonstration théorique *in extenso*, certains proposent des analyses d'entretiens réalisés dans le cadre de recherches de terrain à l'échelon international, avec des experts en finance numérique ou des acteurs de mouvements sociaux.

Puisque tous les domaines que recouvrent les SIC abordent, d'une manière ou d'une autre, la question de la confiance, les chercheurs de la discipline

trouveront un nombre considérable de raisons de lire l'ouvrage, qu'ils travaillent sur la communication des entreprises, de la culture, des savoirs, du politique, des acteurs sociaux, etc. En effet, presque tous nos objets et sujets de recherche questionnent aujourd'hui tout ou partie de signes, d'outils, d'écritures, de médias, de dispositifs, d'organisations sociales, de design, de systèmes symboliques, dans lesquels le numérique et l'internet s'invitent, s'imposent même. D'où l'intérêt pour nous, spécialistes de la communication et de la culture, de prendre au pied de la lettre l'invite de Milad Doueihi et Jacopo Domenicucci à penser de concert la problématique de la confiance et celle des conséquences du numérique sur toutes les formes de communication comme un sujet politique : « La question de la confiance à l'ère du numérique est encore instable, c'est plutôt un sujet politique qu'un sujet historique. Elle est l'enjeu d'une lutte où chacun, par ce qu'il fait ou refuse de faire, participe à une négociation collective. L'enjeu de cette négociation est le visage de la confiance d'aujourd'hui et de demain » (p. 18).

Marie-Sylvie Poli

Avignon Université, LCC, F- 84000 Avignon, France
marie-sylvie.poli[at]univ-avignon.fr

Anaïs THEVIOT, *Big Data électoral. Dis-moi qui tu es, je te dirai pour qui voter*

Lormont, Éd. Le Bord de l'eau, coll. Territoires du politique, 2019, 200 pages

L'ouvrage d'Anaïs Theviot expose les conclusions d'un total de 77 entretiens semi-directifs conduits entre 2012 et 2017 avec des professionnels de la communication numérique des équipes de campagne du Parti socialiste (PS), de l'Union pour un mouvement populaire (UMP) puis des Républicains (LR) et de La République en marche (LREM), ainsi que 20 entretiens avec des directeurs d'organisations de *big data* électoral en France, à savoir : Liegey Muller Pons (LMP, devenue depuis eXplain), Apallian, Emakina, Federavox, Netscouade et DigitalBox. L'objectif est de pénétrer dans les coulisses du *big data* électoral, c'est-à-dire d'interroger les promesses de l'usage des grandes bases de données (*big data*) pour remporter une élection. A. Theviot montre les changements observés à l'issue des campagnes politiques, où le numérique devient central mais non excluant, c'est-à-dire que, en plus des sources des campagnes « classiques » (des données sociodémographiques, des données issues de sondages d'opinion et d'autres retraçant les historiques des votes), s'ajoutent les traces numériques laissées en ligne. Ces dernières semblent être les plus intéressantes

à interroger à la lumière des récents scandales comme celui de l'entreprise Cambridge Analytica, entreprise qui a utilisé des données personnelles, provenant de comptes Facebook, pour influencer la campagne présidentielle de Donald Trump aux États-Unis ainsi que les discussions autour du départ du Royaume-Uni de l'Union européenne (Brexit) et à Donald Trump pendant la campagne présidentielle des États-Unis en 2016. Ce compte rendu expose les principales idées de l'ouvrage tout en respectant son organisation.

Le premier chapitre (p. 15-74) 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'agit 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'autrice identifie à ce moment l'émergence d'une *rhétorique de la proximité* décrite ainsi : « Le candidat s'exprime alors à la première personne 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, des pôles consacrés aux technologies de l'information et de la communication se forment au sein des partis. Par exemple, apparaît « l'adhésion à 20 euros du PS », « 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 questionnaire sur la véritable adhésion de ces partisans. Plus tard, dans le cadre de l'élection présidentielle de 2007, il s'agit de « se saisir du numérique pour militer : [L'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 on le voit avec le site de l'association « Désirs d'avenir » de Ségolène Royal, qui prétend co-construire le programme de la candidate. Néanmoins, A. Theviot souligne le fait 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 traditionnel (volonté de contrôle et formation hiérarchisée) » (p. 38). Indépendamment de cela, l'internet devient un impératif pour les partis.

La chercheuse expose ensuite une série de cas illustrant des mauvaises campagnes sur l'internet tels

les *lip dubs* (vidéos musicales à visée promotionnelle réalisées en plan-séquence et mettant en scène des collaborateurs chantant généralement en *playback*) ou le site participatif de l'UMP « Les Créateurs des possibles » (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 [...] se sent[ai]nt "spammés" par l'envoi massif de mails non désirés » (p. 57). Puis, à partir de 2012, A. Theviot repère l'émergence d'une « rhétorique du changement et de la nouveauté » (p. 70-71), 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. À partir de 2017, même si elles existent depuis toujours, les fausses informations 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'autrice mentionne notamment les exemples d'Alain Juppé avec la caricature « Ali Juppé » (moquant une complaisance supposée du maire de Bordeaux envers l'islam politique) et d'Emmanuel Macron avec la rumeur lui attribuant un compte *offshore*. À l'étranger, les polémiques autour des élections de D. Trump et du Brexit ont entraîné la modification par Facebook de son algorithme pour « privilégier les contenus partagés par les gens "de confiance" dans le fil d'actualité des usagers » (p. 72) et la mise en relation avec des agences de *fact-checking* « dans le but de diminuer les contenus suspects » (p. 73).

Le deuxième chapitre (p. 75-118) interroge la fabrique de données et de logiciels à des fins politiques 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 de Barack Obama (2008) est comprise comme le « point de départ des *data analytics* » (p. 82), ce qui a entraîné une hausse de dépenses dans le domaine des données (collecte et traitement). En France, à la différence de la campagne de B. Obama, « l'usage des *data* en politique reste

très dépendant des budgets » (p. 87-88). Également, le fichage par ethnie ou religion est difficilement applicable en France du fait de la protection des données personnelles. 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). En somme, la complexité de la constitution des bases de données est démontrée par divers cas comme le non-partage des données des primaires du PS en 2012 ou l'essai de l'UMP d'interdire la récolte d'adresses électroniques en 2011.

D'ailleurs, dans ce chapitre, l'autrice s'intéresse à l'usage des logiciels électoraux en France. Ces derniers incluent un premier type de logiciels, comme ceux de LMP (*Cinquante plus un*) et de Federavox, 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, textos, etc.

Dans le prolongement, le troisième chapitre (p. 119-176) aborde la manière dont les campagnes électorales ont intégré les *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, s'inspirant de la campagne de B. Obama en 2008, l'équipe de François Hollande a cherché à cibler en amont « les zones où la proportion d'abstentionnistes de gauche [était] la plus forte, pour rendre encore plus efficace [son] action » (p. 124). La suite consistait à faire une campagne de porte-à-porte efficace en ne dépassant pas les cinq minutes par personne visitée. 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).

Par ailleurs, le marché autour du *big data* électoral se développe. En 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). Sur ce sujet, A. Theviot constate un changement dans 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 » (p. 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. L'autrice prend l'exemple de l'usage de *NationBuilder* 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. Celui-ci a permis de créer un site web, de gérer les médias sociaux, de construire une base de données (comprenant celles des réseaux sociaux non autorisées par la Cnil), d'avoir un *backoffice* du site, de gérer les SMS, la *mailing list* et les messages vocaux et d'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).

Dans sa conclusion, A. Theviot fait allusion à l'intitulé de son ouvrage tout en distinguant l'univers des possibles entre les États-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'envoierai 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 règlement général sur la protection des données (RGPD) entré en vigueur en 2018, le scandale Cambridge Analytica et la méfiance des citoyens envers les Gafam (acronyme désignant les principales entreprises de l'internet : Google, Apple, Facebook, Amazon, Microsoft). Enfin, l'ouvrage présente un travail de récolte de données et d'analyse soigneux, ce qui en fait une référence pour les chercheurs et chercheurs intéressés par la communication politique en ligne.

Camila Pérez Lagos

Université Sorbonne Nouvelle-Paris 3, CIM, F-75005

Paris, France

perez.lagos.camila[at]gmail.com