



Partecipazione e Conflitto
 * *The Open Journal of Sociopolitical Studies*
<http://siba-ese.unisalento.it/index.php/paco>
 ISSN: 1972-7623 (print version)
 ISSN: 2035-6609 (electronic version)
PACO, Issue 11(2) 2018: 528-543
 DOI: 10.1285/i20356609v11i2p528

Published in July 15, 2018

Work licensed under a Creative Commons Attribution-Non commercial-Share alike 3.0 Italian License

SYMPOSIUM/5

HOW DATA-DRIVEN RESEARCH FUELLED THE CAMBRIDGE ANALYTICA CONTROVERSY

Annika Richterich
 Maastricht University

1. Introduction

It is widely agreed that *Cambridge Analytica's* data business practices were, to say the very least, dubious. In early 2018, their "dirty politics" (Milan and van der Velden 2018, 1) made global headlines.¹ Not only *Cambridge Analytica Ltd.* (CAL) and its British parent firm *SCL² Elections Ltd.* (SCL) found themselves in the spotlight: the social networking corporation *Facebook Inc.* and a little-known company called *Global Science Research Ltd.* (GSR) were also under scrutiny for their involvement in the retrieval of personal data from up to 87 million Facebook users.

¹ News outlets had reported on the company earlier already. For instance, the *Guardian* reported on CAL/SCL's ties with Ted Cruz's presidential campaign (Davies 2015) and for *Das Magazin* and *Vice*, Grassegger and Krogerus (2016; 2017) focused on the role of academic research for the company's alleged voter manipulation. Especially their 2017 article attracted quite some attention, but it was also criticised for overstating and mystifying the effectiveness of strategies taken up by CAL (see e.g. Karpf 2017; Baldwin-Philippi 2017). I will come back to this point below.

² SCL stood originally for *Strategic Communication Laboratories*.

In May 2018, CAL and SCL (including multiple, affiliated companies, such as *SCL Analytics Ltd.* and *SCL Commercial Ltd.*) filed for bankruptcy in the United Kingdom (UK); soon afterwards, their US counterparts followed suit. They are being legally investigated in the UK and the US (BBC 2018; Reuters 2018). Facebook CEO Mark Zuckerberg had to testify before US Congress and to justify the corporation's role in exposing users' data in a hearing at the European Parliament. Also, GSR-cofounder and University of Cambridge (UoC) researcher/lecturer Aleksandr Kogan received critical attention.

In 2013, Kogan had developed the quiz app "This Is Your Digital Life", harvesting data from Facebook users who installed the app as well as from their friends. The latter was possible due to Facebook's pre-2014 terms of service (Wong, Lewis and Davies 2018). Initially, the data were reportedly meant to be used in his research.³ Later on though, through GSR, Kogan passed the data on to SCL elections. GSR was co-directed by Kogan and his UoC colleague Joseph Chancellor. In late 2015, Chancellor left GSR to work as quantitative social psychologist and user experience design (UX) researcher for Facebook (Lewis & Wong, 2018). In the same year, Facebook learned that CAL had received its user data from GSR (Constine 2018). Only in early 2018 however, Kogan, the 'This is your digital life' app and CAL/SCL were banned from the social network (Stahl 2018).

In this commentary, I will discuss the controversy around CAL and other actors/companies with regards to two main points. First, the paper examines the societal implications of the scandal: it focuses on the nexus of big data, democratic elections and citizens' perceived freedom of choice as voters. Second, I interrogate the role of GSR in facilitating CAL's immoral data practices. More generally, I reflect on the role of big data-driven research and related news reporting for the promotion of CAL's claims.⁴ I argue that claims concerning the effectiveness of CAL's strategies were reinforced by the narrative that the company drew on scientifically grounded, big data-driven methods and models. I suggest that the involvement of academics in retrieving data and CAL's purportedly scientifically substantiated approach lent credibility to the effectiveness of its services. This also fed into persistent uncertainty regarding the case's implications for individuals' perceived freedom of choice as voters.

The scandal around CAL, SCL, Facebook and GSR is particularly relevant, not only because of the public attention it attracted: on the one hand, it is a notable case, since it

³ In a public statement, the University of Cambridge explained: "In 2015, Dr Kogan applied to the University for ethical approval to use data collected on behalf of GSR for his academic research. His application was reviewed and subsequently rejected" (University of Cambridge 2018).

⁴ While I consider Facebook's involvement just as problematic as the role of CAL/SCL and GSR, this commentary will pay comparatively less attention to the social network. This should however not be seen as reflecting e.g. the responsibility assigned to each of these actors: instead, it is meant to ensure a clear focus of this relatively short paper.

speaks to the relevance of big data *for* politics and political campaigning (see e.g. Gonzalez 2017; Hersh 2015; Tufekci 2014). On the other hand, it concerns the politics *of* big data (see e.g. Noble 2018; Sætnan, Schneider and Green 2018; Coté, Gerbaudo and Pybus 2016)

First, the case emphasises the relevance of big data for contemporary politics, because of the influential narrative that data and analytics were crucial to effective political campaigning and electoral success. The controversy around CAL circles substantially around the issue if and how the company was able to influence individuals' voting behaviour. Second, the scandal highlights the politics of big data by once more illustrating a crucial, though still all too often downplayed and strategically neglected point: data are never neutral – no matter on what scale and how (allegedly) 'unobtrusively' they are retrieved (van Dijck 2014; Gitelman 2013). Data are normative and influential in that they are societally embedded and may be e.g. used to give credibility to claims and arguments, to advocate or undermine certain causes.

This commentary explores these two sides of the same coin and how they are interconnected. I start by examining how the CAL case was discussed with regards to its relevance for contemporary politics. Subsequently, I show that the predominant, more general idea of big data's alleged superiority, notably in combination with their scientific use, played a significant role too.

2. A wake-up call?

What was and is at stake in debates concerning CAL, SCL and related actors/corporate entities is not 'only' the privacy of these users: the scandal also raises questions regarding individuals' autonomy and freedom of choice as voters. The data analytics firm is not merely controversial, because of how it retrieved the data – despite their approaches being highly problematic as such. Its data practices led to public outcry, because of what the company claimed and was said to achieve: the nudging and manipulation of individuals, with the help of (social media) microtargeting. By supposedly profiling users based on, for example, their Facebook 'Likes' or preferences and then placing targeted, political advertising, the company is reputed to have swayed opinions and voting behaviour. This applies, for example, to the 2016 UK referendum on the country's European Union (EU) membership and to Donald Trump's 2016 US presidential campaign (Cadwalladr and Graham-Harrison 2018; Rosenberg, Confessore and Cadwalladr 2018; Reuters Technology News 2018).

The news reporting and public debates interrogated if and how CAL might have affected previous elections (see also Gonzalez, 2017). Moreover, it was discussed how the revelations were relevant to upcoming polls. In early 2018, speaking at a symposium in Vienna, Věra Jourová, *EU Commissioner for Justice, Consumers and Gender Equality*, described the CAL scandal as a “wake-up call” that “[...] raises serious concerns about our collective freedom as voters and politicians.” (2018) She furthermore argued:

[W]e have to understand if these practices might have had relevance for elections or referenda in Europe. If only one country's elections are at risk of being manipulated, this has an impact on our whole Union. And this is a big concern, in particular ahead of the upcoming [May 2019, added by the author] European Parliament elections. (Jourová 2018)

As also indicated in this quote, while CAL's data retrieval practices have been widely condemned as immoral and potentially illegal, it is still contested and under examination to what extent their attempts were indeed effective. This uncertainty does not only alarm politicians such as Jourová; it also adds to concerns and insecurity among potential voters. Wetherell (2018) gets to the heart of the issue when observing that: “The recent revelations about Cambridge Analytica have prompted a new and strange kind of anxiety: What if your political opinions are not actually yours?” (Wetherell 2018).

One may argue that this question is not necessarily new: it goes, at least, as far back as ancient Greek deliberations on rhetoric, i.e. persuasion (Jones and Simons, 2011, chapter 1). Plato criticised rhetoric as inevitably being devoid of any educational value, merely deceiving the audience. In contrast, Aristotle defended rhetoric as double-edged sword, also vital to the imparting of knowledge and truth. Echoing the latter sentiment, though centuries later, propaganda was described as a tool “[...] to be used either for good, bad, or amoral goals – similar to advertising or public relations. The primary goal is *mass persuasion* [...].” (Seidman 2008, 7) Long before recent efforts in using big data-driven techniques for political campaigning, technology has been a crucial factor for the changing face of persuasion and propaganda. Advances in printing processes (e.g. the use of steel printing plates) allowed for the inclusion of images, such as engravings of candidates; the popularisation of photography enabled the use of photographic images in the 1920s and digital production conditions eased their manipulation in the 1990s (Seidman 2008, 9–12).

As Wetherell (2018) suggests: “Theorists from various different disciplines have long known that people's desires and preferences are shaped by massive forces beyond

their control.” However, now it is not merely rhetoric that individuals may fear, but already the very selection of digital content that they may or may not encounter. This concern is closely related to recent developments in political campaigning and attempts at increasingly data-driven (micro)targeting strategies to address and mobilise voters (Hersh 2015).

A main anxiety that is instilled in voters and politicians alike is that even efforts aimed at deconstructing political messages could be largely futile: because such messages are tailored to e.g. their fears or interests; and because voters’ very chances for encountering certain kinds of content may be tilted. Wetherell (2018) suggests that, while triggered by the CAL scandal, this anxiety is rooted in an “[...] individualized way of thinking about voting that emerged in the nineteenth century”. This commentary will be less concerned with the historical roots of the issue though. Instead, it interrogates why the uncertainty regarding the effectiveness of CAL’s data practices is problematic. More specifically, I focus on the case’s significance for democratic elections.

CAL staff and affiliates portrayed themselves as powerful influencers and victims alike – depending on which strategy they saw fit for the respective audience. Shortly after Donald Trump won the US presidential election in November 2016, the company released a (self-)congratulatory press release. It included a quote by CAL CEO Alexander Nix, stating: “We are thrilled that our revolutionary approach to data-driven communications played such an integral part in President-elect Donald Trump’s extraordinary win [...]. It demonstrates the huge impact that the right blend of cutting-edge data science, new technologies, and sophisticated communication strategies can have.” (Cambridge Analytica 2016) In a pitch to an undercover *Channel 4 News* reporter, disguised as potential client, company representatives boasted about CAL’s manipulative, political campaigning strategies and their impact on elections globally (Politico 2018; see also Channel 4 News 2018; Cadwalladr 2018). The video report, including recordings from meetings with CAL staff, first aired in March 2018. However, since such claims are simply part of the company’s marketing sales efforts, they do not say much about their actual effectiveness. The article title “Shadowy puppet masters or snake oil salesmen?” (Sutton 2018) pointedly summarises the discussion that unfolded afterwards.

The uncertainty whether or not CAL’s data practices were as effective as i.a. the company originally claimed was not least rooted in a longstanding discourse of ‘big data’ as the “holy grail of behavioral knowledge” (van Dijck 2014, 199). ‘Big data’ is a broad and hyped umbrella term for digital data retrieved from various sources on a large scale. The term is commonly associated with the ‘three Vs’: standing for (high) volume, variety and velocity – as main characteristics. According to van Dijck (2014) the

big data phenomenon has facilitated the ideology of *dataism* “[...] as a belief in a new gold standard of knowledge about human behaviour” (p. 198). This applies also to scientific research: such research may draw on data which are e.g. retrieved from social networking sites (SNS). In doing so, it also validates big data practices – thanks to public trust in science (202ff.; see also Richterich 2018). Unfortunately, as Van Dijck argues, i.a. “[i]nformation scientists often uncritically adopt the assumptions and ideological viewpoints put forward by SNSs and data firms” (2014, 201). This may lead to problematic entanglements between scientific and corporate practices – as also illustrated in the case of the CAL scandal.

Due to the spreading ideology of dataism and increasingly complex interrelations between scientific and tech-commercial objectives, big data-driven practices are highly persuasive. This is also the case where their actual effectiveness may not be explained/justified in detail and hence questionable. Terms such as ‘big data’ and ‘data analytics’ have become potent buzzwords, relying heavily on dataism and an often unquestioned trust in the power of (big) data. Such dataism also plays into the hands of actors such as CAL, since it gives credibility to their marketing efforts and service sales pitches.

Nevertheless, a few months after the *Channel 4 News* video was published, during an EP hearing in June 2018, Alexander Nix described himself as “victimised”, blaming in particular liberal media and former employee/‘whistleblower’ Christopher Wylie for the company’s downfall. It should be noted here that many of his comments were directly refuted during the hearing, e.g. by “[...] Brendan O’Hara, a member of the Scottish National Party, [who] told Nix during the hearing. ‘By no stretch of the imagination can you be seen as the victim.’” (Lapowsky 2018) CAL obstructed insights into its business practices, e.g. by deleting previously held Facebook data and then denying having ever received these data (Waterson 2018). Since the unfolding of the scandal, company representatives have been ambiguous and obscure about their big data-driven practices and their effectiveness – in stark contrast to earlier claims targeted at potential clients.

This obstruction of insight into the case may have ultimately contributed to CAL’s downfall. However, their ambiguity also feeds into uncertainty regarding how such companies, with the help of social media data and analytics, may (aim to) influence individuals’ decisions as voters. This also risks instilling abovementioned anxieties, by casting doubt on individuals’ perceived autonomy and freedom of choice when voting. While many have argued that the influence of CAL is highly doubtful (Martínez 2018; Wetherell, 2018; Trump 2018; see also Sutton 2018), this risk is also one of the reasons why (ongoing) investigations are urgently needed to safeguard public trust in the signif-

icance of elections. Jourová's description of the case as "wake-up call" referred mainly to the responsibility of politicians to ensure clarification and eventually needed counter-measures. Nevertheless, one may also ask what the controversy may mean to voters. A main concern here is that the debates around CAL run the risk of de-motivating and frustrating i.a. European voters – which appears particularly problematic on EU level, given the steady decrease in voter turnout in European elections (European Parliament 2014).

In this section, I have argued that much uncertainty remains with regards to the 2018 controversy concerning CAL/SCL, Facebook and GSR. This development risks fostering anxiety of potential voters, with regards to their perceived autonomy and freedom of choice in elections. So far, I have mainly explored the role of CAL/SCL in this process. In the following section though, I will highlight another aspect of the scandal: the involvement of the initially mentioned scientist Aleksandr Kogan. I highlight his role, because it has been long argued that "[un]critical acceptance of datafication's underpinning ideological and commercial premises may well undermine the integrity of academic research in the long run" (van Dijck 2014, 206). Nevertheless, it still seems to take a scandal to bring such research up for discussion. Moreover, I look at the relevance of news reporting that associated CAL's approach with scientific, data-driven methods.

3. The role of academic, big data-driven research

As mentioned in the introduction, CAL/SCL received Facebook data from University of Cambridge (UoC) lecturer/researcher Aleksandr Kogan. In 2014, Kogan developed the Facebook personality-quiz app 'This is your digital life'. With the help of this app, he retrieved data from Facebook users who installed the app and from their friends. Kogan's work was directly inspired by a personality app previously developed and explored by some UoC colleagues (see also Vaidhyanathan 2018, chapter 6). The Facebook personality quiz app 'MyPersonality' was developed by David Stillwell in 2007, while studying at University of Nottingham. Later, he started working as lecturer at UoC. In 2009, Michal Kosinski joined the project, as PhD student at UoC (see also Kosinski n.d.) Both were part of UoC's *The Psychometrics Centre* which is "[...] dedicated to research, teaching and product development in both pure and applied psychological assessment." (The Psychometrics Centre n.d.) Psychometric research focuses largely on tools and theories for psychological measurements, i.e. testing and measuring (for example) attitudes, personality traits or skills (see also Schoenherr and Hamstra 2016).

According to the app/research project website (Kosinski n.d.), 6 million people participated in the app-based 'MyPersonality' questionnaires: users were able to decide whether they would like to donate their scores and data for research purposes, indicating their consent through an opt-in statement; about 40% agreed to this (see also Kosinski n.d.; Matz, Kosinski, Nave, and Stillwell 2017). In 2014, Stillwell and Kosinski were approached by Aleksandr Kogan, on behalf of SCL. Kogan inquired if the researchers would provide SCL with the 'MyPersonality Project' database and model. The latter was partly based on the OCEAN scale for the 'big five' personality traits, i.e. openness, conscientiousness, extroversion, agreeableness, neuroticism (Vaidhyanathan 2018, chapter 6). However, the collaboration did eventually not emerge. The reasons for this are contested: Kogan stated that there was disagreement about the payment that Stillwell and Kosinski allegedly requested from SCL (Lewis, Grierson and Weaver 2018). Other sources however highlight that Stillwell and Kosinski had ethical concerns regarding SCL's data practices and objectives (Grassegger and Krogerus 2017).

Since Kogan was thus not able to receive the data and model from Stillwell and Kosinski, he and GSR co-founder Joseph Chancellor proceeded on their own. Through their 'This is your digital live' personality quiz app, they retrieved the needed data. While still working at UoC, these data were passed on to SCL through the newly founded company GSR. Within their department, Kogan and Stillwell were confronted with and criticised for their approach, being i.a. accused of trying to realise a "'get rich quick' scheme" (Lewis, Grierson and Weaver 2018). Kogan insisted that he never received a salary from SCL, but he did concede that SCL paid GSR £230,000 at some point (Lomas 2018).

Already in September 2016, at the *Concordia Summit*, CAL CEO Nix talked about how the company had harnessed "the power of Big Data and psychographics in the electoral process" (Grassegger and Krogerus 2017). Nix claimed that by combining behavioural science, notably the Big Five/OCEAN Model, big data analytics and targeted political advertising messages CAL was able to steer election results. While he mentioned neither Kosinski and Stillwell's nor Kogan and Chancellor's work, parallels between the approaches were noted not only by Kosinski, but also by journalists (Grassegger and Krogerus 2017). In Grassegger and Krogerus' widely circulated article "The Data That Turned the World Upside Down" (2017), Kosinski's distances himself explicitly from the approaches taken by CAL and GSR. At the same time however, the article contributes to the idea that CAL's strategies were modelled on scientific methods.⁵

⁵ The authors do raise the issue: "But to what extent did psychometric methods influence the outcome of the election? When asked, Cambridge Analytica was unwilling to provide any proof of the effectiveness of

This idea in turn lends credibility to the effectiveness of the approaches (as strategically promoted by companies such as CAL), despite this impact being questionable. In this sense, also Baldwin-Philippi (2017) argues: “[w]hile much has been written about the possibilities of data driven campaigning, the on-the-ground realities are often much less precise and much less novel than journalistic coverage implies” (627). The risk thus lies in transferring the credibility associated with scientific methods to ethically dubious, political data business practices. I thus argue that the discursive establishment of this myth, substantiated with reference to the use of scientific, data-driven methods, is one of the main reasons why it has been so difficult to relativize CAL’s effectiveness. As explained in the previous section, I have moreover argued that the uncertainties regarding CAL’s approaches are potentially damaging for individuals’ assessment of their own freedom of choice and autonomy as voters. In consequence, debunking the myth of CAL as rooted in scientific methods is a vital factor to address the significance of elections in a big data era.

It is of course little surprising that Kosinski and others have asserted the effectiveness of their methods and the value of psychometrics. This position is needed to uphold the relevance of their own, academic research. Yet, in contributing to news reporting that links this research to CAL’s business practices, they simultaneously lent credibility to the effectiveness of the company’s services. Still, as also Karpf (2017) points out, “[...] there is no evidence that Cambridge Analytica has solved the practical challenges of applying psychometrics to voter behaviour”. Thus, the researchers entered complex, discursive interdependency linking the credibility of their own research to the purported effectiveness of CAL. What is therefore urgently needed are reports on CAL that separate its marketing and strategic communication, claimed to be grounded in scientific methods, from its actual practices. Needless to say though, so far it has been almost impossible to receive any reliable insights into the latter from those involved.

Of all people entangled in the scandal, Kogan has notably questioned and downplayed the effectiveness of CAL’s strategies. In a BBC radio interview (cited in *The Guardian*), he stated:

I personally don’t think micro-targeting is an effective way to use such data sets. It could have only hurt the campaign. What Cambridge Analytica has tried to sell is magic. And it made claims that this is incredibly accurate and it tells you everything

its campaign. And it is quite possible that the question is impossible to answer.” (Grassegger and Krogerus 2017) Yet, they proceed with a section on “clues” for the effectiveness.

there is to tell about you, but the reality is that it's not that. If you really work through the statistics ... those claims quickly fall apart. (Weaver 2018)

Alas, such statements carry only little persuasive power, as they are easily dismissed as part of Kogan's efforts in justifying his own role and the implications for democratic elections. They are closely related to how much responsibility he needs to assume for sharing the data with SCL. In consequence, such assurances are unlikely to make a convincing case for the ineffectiveness of CAL's approach.

Data-driven research, located at leading universities, thus played a major role in fuelling uncertainty regarding CAL's impact on recent elections. On the one hand, this is linked to the direct involvement of academics such as Kogan and Chancellor in collecting data that were later commercially provided to SCL/CAL, via their company *Global Science Research* (GSR). On the other hand, this is related to news reporting that described CAL's approaches as rooted in scientific methods. The emerging myth that CAL was relying on data-driven, scientifically inspired and thus allegedly effective methods is damaging and problematic insofar as it calls voters' freedom of choice into question. It revives a debate that one can historically, recurrently witness with the emergence of new technologies: while e.g. mass media triggered discussions on mass persuasion and 'manufactured consent' (Herman and Chomsky, 1988; see also Thomson, 1977), social media have prompted concerns that individuals may receive ideologically infused messages, tailored to their personal fears, interests and preconceptions. This time round though not psychoanalysis, interviews and focus groups (see e.g. Packard 1957)⁶, but psychometrics and social media data are described as key to individuals' decision making.

4. Conclusion

This commentary started from the observation that the controversy regarding CAL/SCL's dubious data business practices instigated uncertainty among voters and politicians alike. CAL's use of Facebook user data and its allegedly effective microtargeting of voters, to influence election results, called individuals' autonomy and freedom of choice as voters into question. While the company's data retrieval and marketing strategies are widely condemned as immoral, it is however debatable whether its methods were indeed also effective (Karpf 2017; Gonzalez 2017).

⁶ Despite having been a popular, influential book, it received "[...] widespread academic and advertising industry criticism, in part for its sensationalist, unsubstantiated writing [...]" (Nelson 2008, 113).

Ongoing debates around these issues speak to the relevance of big data *for* politics, on the one hand, and the politics *of* big data, on the other hand. Big data and related analytics have been promoted as part and parcel of effective, contemporary political campaigning. The scandal around CAL demonstrates however that such big data-driven strategies may clash with what is morally appropriate and, potentially, legal. While calling attention to CAL's immoral data collection methods, among other things, the scandal also initiated discussions on if and how the company was able to influence individuals' voting behaviour.

The latter point also hints at the politics of big data: the controversy around this question is related to the predominant idea of big data's alleged superiority, notably in combination with their scientific use. Big data have been widely pitched and hyped as powerful backbone of scientific research – and this image was significant for CAL's business claims as well as the uncertainty following the scandal. In addition to the significance of widespread *dataism* (van Dijck 2014), the involvement of academics and news reporting on the inspirational role of big data-driven research further fuelled the Cambridge Analytica controversy.

It has been proposed that CAL's claims were little more than “snake oil” (see e.g. Sutton 2018), aimed at selling questionable services. Yet, its questionable claims were substantiated by the idea that the company's approaches were modelled on scientific, big data-driven methods and relying on data collected by academics. This paper therefore examined how the involvement of academics in retrieving data and CAL's purportedly scientifically inspired and grounded approach lent credibility to the effectiveness of its services. I argued that the uncertainty regarding whether CAL's data business practices affected recent elections was linked to discourses in which the company associated itself with – and was associated with – scientific, big data-driven research.

Ongoing investigations into the case as well as balanced news reporting are hence much needed to counter insecurity on side of voters and politicians alike. Countering the myth that CAL's dubious strategies were built upon scientific, data-driven models and methods may help relativize emerging anxiety concerning voters' (lack of) autonomy. Having said that, this perspective should not distract from CAL/SCL's immoral intentions and business practices – which may have moreover been illegal. At the same time, one may hope that the controversy also functions as a reminder of the importance of vigilant, critical and media literate citizenship in the digital age.

References

Baldwin-Philippi J. (2017), “The Myths of Data-Driven Campaigning”, *Political Communication* 34(4): 627-633.

- BBC (2018, May 3), "Cambridge Analytica: Closure 'will not stop investigation'", Retrieved from: <https://www.bbc.com/news/uk-43985186>.
- Cadwalladr C. (2018, March 18), "'I made Steve Bannon's psychological warfare tool': meet the data war whistleblower", *The Guardian*, Retrieved from <https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump>.
- Cadwalladr C., E. Graham-Harrison (2018, March 17), "Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach", *The Guardian*. Retrieved from <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
- Cambridge Analytica (2016, November 9), "Cambridge Analytica Congratulates President-elect Donald Trump and Vice President-elect Mike Pence", [press release], PR News Wire, Retrieved from <https://www.prnewswire.com/news-releases/cambridge-analytica-congratulates-president-elect-donald-trump-and-vice-president-elect-mike-pence-300359987.html>.
- Channel 4 News (2018), "Cambridge Analytica Uncovered: Secret filming reveals election tricks", *YouTube*. Retrieved from www.youtube.com/watch?time_continue=166&v=mpbeOCKZFfQ.
- Constine J. (2018, April 10), "Zuckerberg admits it was a mistake not to ban Cambridge Analytica's ads", *TechCrunch*, Retrieved from <https://techcrunch.com/2018/04/10/ban-cambridge-analytica/?guccounter=1>.
- Coté M., P. Gerbaudo, J. Pybus (eds.), (2016), "Politics of Big Data", *Digital Culture and Society*, 2(2).
- Davies H. (2015, December 11), "Ted Cruz using firm that harvested data on millions of unwitting Facebook users", *The Guardian*, Retrieved from <https://www.theguardian.com/us-news/2015/dec/11/senator-ted-cruz-president-campaign-facebook-user-data>.
- European Parliament [TNS/ScytI in cooperation with the European Parliament] (2014), Results of the 2014 European elections, Retrieved from <http://www.europarl.europa.eu/elections2014-results/en/turnout.html>.
- Gitelman L. (ed.), (2013), *Raw data is an oxymoron*, Cambridge, MIT Press.
- Gonzalez R. J. (2017), "Hacking the citizenry?: Personality profiling, 'big data' and the election of Donald Trump", *Anthropology Today*, 33(3): 9-12.
- Grassegger H., M. Krogerus (2016, December 3), „Ich habe nur gezeigt, dass es die Bombe gibt“, *Das Magazin*, Retrieved from <https://www.dasmagazin.ch/2016/12/03/ich-habe-nur-gezeigt-dass-es-die-bombe-gibt/?reduced=true>.

- Grassegger H., M. Krogerus (2017, January 28), The Data That Turned the World Upside Down, *Vice*, Retrieved from https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win.
- Herman, E., S., N. Chomsky (1988), *Manufacturing consent: The political economy of the mass media*, London: Vintage.
- Hersh E. D. (2015), *Hacking the electorate: How campaigns perceive voters*, Cambridge: Cambridge University Press.
- Jones J.G., H.W. Simons (2011), *Persuasion in Society*, London, Routledge.
- Jourová V. (2018, June 7), "Speech by Commissioner Jourová on media and democracy in Europe at the Media Symposium", Vienna, Retrieved from http://europa.eu/rapid/press-release_STATEMENT-18-4095_en.htm.
- Karpf D. (2017), "Will the Real Psychometric Targeters Please Stand Up?", *Civicist*, Retrieved from <https://civichall.org/civicist/will-the-real-psychometric-targeters-please-stand-up>.
- Kosinski M. (n.d.), "My personality project", Retrieved from <https://sites.google.com/michalkosinski.com/mypersonality>.
- Lapowsky I. (2018, June 6), "Former Cambridge Analytica CEO Faces His Ghosts in Parliament", *Wired*, Retrieved from <https://www.wired.com/story/alexander-nix-parliament-testimony-cambridge-analytica>.
- Lewis P., J. Grierson, M. Weaver (2018, March 24), "Cambridge Analytica academic's work upset university colleagues", *The Guardian*, Retrieved from <https://www.theguardian.com/education/2018/mar/24/cambridge-analytica-academics-work-upset-university-colleagues>.
- Lomas N. (2018, April 24), "Kogan: 'I don't think Facebook has a developer policy that is valid'", *TechCrunch*, Retrieved from <https://techcrunch.com/2018/04/24/kogan-i-dont-think-facebook-has-a-developer-policy-that-is-valid/?guccounter=1>.
- Martínez A.G. (2018, March 19), "The Noisy Fallacies of Psychographic Targeting", *Wired*, Retrieved from <https://www.wired.com/story/the-noisy-fallacies-of-psychographic-targeting>.
- Matz S. C., M. Kosinski, G. Nave, G. D. J. Stillwell (2017), Psychological targeting as an effective approach to digital mass persuasion, *Proceedings of the National Academy of Sciences*, 114(48): 12714-12719.
- Milan S., L. van der Velden (2018), "Reversing Data Politics: An Introduction to the Special Issue", *KRISIS. Journal for contemporary philosophy*, Retrieved from <http://krisis.eu/reversing-data-politics-an-introduction-to-the-special-issue>.

- Nelson M.R. (2008), "The Hidden Persuaders: Then and Now", *Journal of Advertising*, 37(1): 113-126.
- Noble S. U. (2018), *Algorithms of Oppression: How search engines reinforce racism*, New York: NYU Press.
- Packard V. (1957), *The hidden persuaders*, London: Pelican.
- Politico, (2018, March 19), Cambridge Analytica boasts of dirty tricks to swing elections, Retrieved from <https://www.politico.eu/article/cambridge-analytica-boasts-of-dirty-tricks-to-swing-elections>.
- Reuters (2018, May 16), "U.S. investigating Cambridge Analytica: New York Times", Retrieved from: <https://www.reuters.com/article/us-facebook-privacy-cambridge-analytica/u-s-investigating-cambridge-analytica-new-york-times-idUSKCN1IH02J>.
- Reuters Technology News (2018, April 16), "Lawmakers publish evidence that Cambridge Analytica work helped Brexit group", Retrieved from <https://www.reuters.com/article/us-facebook-cambridge-analytica-britain/lawmakers-publish-evidence-that-cambridge-analytica-work-helped-brexit-group-idUSKBN1HN2H5>
- Richterich A. (2018), *The big data agenda: Data ethics and critical data studies*, London, University of Westminster Press.
- Rosenberg M., N. Confessore, C. Cadwalladr (2018, March 17), How Trump Consultants Exploited the Facebook Data of Millions, *The New York Times*, Retrieved from <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html>.
- Sætnan A. R., I. Schneider, N. Green (2018), *The Politics and Policies of Big Data: Big Data, Big Brother?*, London: Routledge.
- Schoenherr J. R., S. J. Hamstra (2016), Psychometrics and its discontents: An historical perspective on the discourse of the measurement tradition, *Advances in Health Sciences Education*, 21(3):719-729.
- Stahl L. (2018, April 22), "Aleksandr Kogan: The link between Cambridge Analytica and Facebook", *CBS News*, Retrieved from <https://www.cbsnews.com/news/aleksandr-kogan-the-link-between-cambridge-analytica-and-facebook>.
- Sutton J. (2018). Shadowy puppet masters or snake oil salesmen? *The Psychologist*, 31. Retrieved from <https://thepsychologist.bps.org.uk/volume-31/may-2018/shadowy-puppet-masters-or-snake-oil-salesmen>.
- The Psychometrics Centre (n.d.), "About us", Retrieved from <https://www.psychometrics.cam.ac.uk/about-us>.

- Thomson O. (1977), *Mass persuasion in history: an historical analysis of the development of propaganda techniques*, New York: Crane, Russak & Co.
- Trump K. S. (2018, March 23), "Four and a half reasons not to worry that Cambridge Analytica skewed the 2016 election", *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/monkey-cage/wp/2018/03/23/four-and-a-half-reasons-not-to-worry-that-cambridge-analytica-skewed-the-2016-election/?noredirect=on&utm_term=.2739d3961520.
- Tufekci Z. (2014), "Engineering the public: Big data, surveillance and computational politics", *First Monday*, 19(7), Retrieved from <http://firstmonday.org/ojs/index.php/fm/article/view/4901/4097>.
- University of Cambridge (2018, March 23), "Statement from the University of Cambridge about Dr Aleksandr Kogan", Retrieved from <https://www.cam.ac.uk/notices/news/statement-from-the-university-of-cambridge-about-dr-aleksandr-kogan>.
- Vaidhyanathan S. (2018), *Antisocial Media: How Facebook Disconnects Us and Undermines Democracy*, Oxford: Oxford University Press.
- Van Dijck J. (2014), "Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology", *Surveillance & Society*, 12(2):197-208.
- Waterson J. (2018, June 6), "Cambridge Analytica's 'victimised' ex-chief lambasts liberal media", *The Guardian*, Retrieved from <https://www.theguardian.com/uk-news/2018/jun/06/cambridge-analytica-destroyed-by-bitter-and-jealous-whistleblower>.
- Weaver M. (2018, March 21), "Facebook scandal: I am being used as scapegoat – academic who mined data", *The Guardian*, Retrieved from <https://www.theguardian.com/uk-news/2018/mar/21/facebook-row-i-am-being-used-as-scapegoat-says-academic-aleksandr-kogan-cambridge-analytica>.
- Wetherell S. (2018, April 3), "Cambridge Analytica and the Death of the Voter", *Verso Books*, Retrieved from <https://www.versobooks.com/blogs/3720-cambridge-analytica-and-the-death-of-the-voter>.
- Wong J., Lewis P., Davies H. (2018, April 24), "How academic at centre of Facebook scandal tried – and failed – to spin personal data into gold", *The Guardian*, Retrieved from <https://www.theguardian.com/news/2018/apr/24/aleksandr-kogan-cambridge-analytica-facebook-data-business-ventures>.

Author's Information

Annika Richterich (PhD) is an assistant professor in Digital Culture at Maastricht University's Faculty of Arts and Social Sciences (NL). She studied sociology at University of Auckland (NZ) as well as media studies and economics at University of Siegen (GER). Her research has been published in international, peer reviewed journals such as *Convergence and Information, Communication & Society*. In early 2018, she published her open access book *The Big Data Agenda: Data Ethics and Critical Data Studies* with University of Westminster Press.