

Operationalizing data-driven campaigning: designing a new tool for mapping and guiding regulatory intervention

Rachel Gibson^a, Esmeralda Bon^a and Andrea Römmele^b

^aUniversity of Manchester, Manchester, UK; ^bHertie School of Governance, Berlin, Germany

ABSTRACT

Since the *Cambridge Analytica* scandal, governments are increasingly concerned about the way in which citizens' personal data are collected, processed and used during election campaigns. To develop the appropriate tools for monitoring and controlling this new mode of "data-driven campaigning" (DDC) regulators require a clear understanding of the practices involved. This paper provides a first step toward that goal by proposing a new organizational and process-centred operational definition of DDC from which we derive a set of empirical indicators. The indicators are applied to the policy environment of a leading government in this domain – the European Union (EU) – to generate a descriptive "heat map" of current regulatory activity toward DDC. Based on the results of this exercise, we argue that regulation is likely to intensify on existing practices and extend to cover current "cold spots". Drawing on models of internet governance, we argue that this expansion is likely to occur in one of two ways. A "kaleidoscopic" approach, in which current legislation extends to absorb DDC practices and a more "designed" approach that involves more active intervention by elites, and ultimately the generation of a new regulatory regime.

ARTICLE HISTORY

Received 19 September 2022
Accepted 12 September 2023

KEYWORDS

Digital; data-driven; campaigning; regulation; online political advertising; GDPR; Digital Services Act; EU

Introduction

The idea that political campaigns are entering a new "fourth" era has gained ground in recent years, particularly among communication scholars (Blumler 2016; Esser and Pfetsch 2020). At the core of this transition is the digitization of society and the growth of what is seen as a new and more technologically intensive "data-driven" form of campaigning (Römmele and Gibson 2020). Work to date on data-driven campaigning (DDC) has focused on a range of important questions including its effect on voter behaviour and attitudes (Endres and Kelly 2018; Kruikemeier, Sezgin, and Boerman 2016; Lewandowsky et al. 2020; Nickerson and Rogers 2014; Turow et al. 2012), patterns of adoption within and across party systems (Anstead 2017; Bennett 2016; Kefford et al. 2022; Kruschinski and Haller 2017), and broader normative consequences for democracy and fair elections

CONTACT Rachel Gibson  rachel.gibson@manchester.ac.uk

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

(Borgesius et al. 2018; Gorton 2016; Persily 2017). In this paper, we shift and re-focus the lens both inward and upward to examine the core practices that constitute DDC and the wider regulatory arena that governs it, with special attention paid to the legislative environment of the European Union. To do so we develop a new organizational and process-centred definition of DDC and derive a set of empirical indicators that allow us to map the current “hot” and “cold” spots in contemporary regulatory approaches across several major democracies.

Through this analysis, we make four main contributions to the study of DDC. First, we add more conceptual clarity and empirical specificity to the current understanding of this new campaign mode by providing a new organizationally based operational definition. Second, we provide an overview of the legislation surrounding DDC within a major governmental context, and insight into the extent to which it forms, or is beginning to form a new and coherent body of campaign regulation or remains fragmented across different authorities and legal frameworks. While there has been some attention paid to regulation of DDC to date, much of it has focused on controls over particular practices such as micro-targeting or big tech companies uses of citizens personal data in the political process (Dobber, Fathaigh, and Zuiderveen Borgesius 2019; Harker 2020; McEvoy 2019; Witzleb, Paterson, and Richardson 2019). Third, we use our findings about regulatory hot and cold spots within the case of the EU to develop two alternative versions of a future regulatory regime. Finally, by specifying empirical indicators of DDC, we provide a basis for future research to study adoption by parties in a more systematic manner, thereby opening up the potential for more comparative analysis of this phenomenon.

Defining and operationalizing data-driven campaigning (DDC) at the party level

The core purpose of this paper is to develop a new operational definition of DDC that can be used to map the extent of control offered by current national and international legislative frameworks over this new form of campaigning. While our focus is on parties as the reference or “default” political organizations most likely to adopt DDC, the definition is not exclusive to them and can be applied to the organizations surrounding individual candidates.

To develop an operational measure of DDC at the organizational or party-actor level, it is important to have a clear understanding of the concept in question. As a recent systematic review by Dommett, Barclay, and Gibson (2023) reveals, however, such definitional precision is lacking in the literature. A number of scholars take an explicitly critical and normative approach, focusing on the threats posed by the new forms of digital campaigning to citizen privacy, informed decision-making and vote manipulation (Borgesius et al. 2018; Gorton 2016; Persily 2017). Other accounts refer in more neutral terms to changes in the tools and techniques being used, typically citing the growing use of “big data” and new computational methods and adoption of particular practices such as micro-targeting or voter propensity scoring (Bennett 2016; Kruikemeier, Sezgin, and Boerman 2016; Nickerson and Rogers 2014). Pinpointing when DDC first emerged is also something of a moving target, although the U.S. Presidential elections of 2008 and 2012 appear to be commonly accepted as key events in the shift into this new “technology intensive” mode of campaigning (Kreiss 2016; Stromer-Galley 2014).

Since the election of Donald Trump in 2016 explicit references to “data-driven campaigning” has increased sharply although consensus on a “joined up” or singular definition remains elusive. Baldwin-Philippi (2017) made one of the first attempts, arguing that DDC has “two main features: targeting, or deciding which messages go to what potential voters at what time during the campaign, and testing, or empirically measuring how well messages perform against one another and using that information to drive content production and further targeting”. Munroe and Munroe (2018) extended this definition to include a broader range of organizational features. According to their “conceptual framework of DDC”, its presence requires that a campaign perceives data as a resource “in its own right on par with volunteer time or money” and “the campaign leadership make decisions about what to do, based on analysis of data ... rather than by instinct, guesswork, intuition, tradition or rules of thumb” (148). This distinction between the practical and more perceptual and structural components of DDC is important and echoes Kreiss (2016) earlier work tracing the rise of “prototype politics” among U.S. parties, and particularly the Democrats, from the early twenty first century. For Kreiss, the emergence of this new mode of electioneering was dependent on both an extensive technical infrastructure and a conducive “networked” organizational culture that allowed for importation and experimentation with innovative ideas. Subsequent comparative work has extended the lens to examine how the understanding and practice of DDC differ across national contexts and has further questioned the idea that a “one size fits all” definition of DDC exists (Kefford et al. 2022). For Dommett, Barclay, and Gibson (2023), such variation suggests that DDC should be defined in more flexible and disaggregated terms that permits its adaptation to local and temporal circumstances.

For the purposes of this paper, we draw on these insights to develop an understanding of DDC that takes into account its technical and organizational characteristics, and locates those features within a broader thematic and flexible framework that can be applied across time and space. In broad terms, we consider DDC distinguishable from earlier modes of electronic and computationally based campaigning in three main regards. First, the size and granularity of data used is much greater, making personalized or micro-messaging on a mass scale possible for the first time. Second, the data used are wholly or predominantly digitally sourced or digitized in nature and their analysis typically involves some degree of statistical and/or algorithmic modelling. Finally, and perhaps most critically, the data, and particularly the results it yields are highly influential to the point of being deterministic in the key decisions of the campaign. From this general conceptual understanding of DDC, we draw out a more specific “process”-based understanding that hinges on the performance of four key stages or genres of activity. For each stage, we describe its overall purpose and the core activities that comprise it. Using this process-based model of DDC, we then present an operational definition of DDC and draw out a series of empirical indicators.

In following this sequential logic, we fulfil the first three of the four stages of concept measurement set out by Adcock and Collier (2001) that involves moving from a background to a systematized concept, and then indicator development. The fourth stage – the deriving of numerical scores for a particular indicator – is something we stop short of here, although the construction and empirical application of such an index to a range of regulatory regimes is clearly a next step in this process. For the purposes of this paper, reaching this penultimate point of indicator generation is sufficient for us

to meet our central goal of mapping which elements of DDC are (and are not) currently subject to regulatory scrutiny.

Operationalizing DDC: four stages of delivery

Based on the understanding of DDC as both a generic and more specifically or contextually defined process that is underpinned by both technical and organizational resources, we draw on the extant literature to logically extrapolate four main stages that can be considered as essential to the delivery of DDC. Each stage is defined by a combination of the technical and human or organizational resources required to support it. Given the focus of the paper, we are particularly interested in identifying its more tangible features that would be visible to governing authorities and thus subject to potential regulation. Thus for each stage, we provide examples of a range of more particular tools, activities and personnel associated with its delivery. Those resources may be internal to the organization, i.e. purpose-built hard and software or “in-house” staff and volunteers that are specifically recruited to carry out key tasks or externally provided on a consultancy and commercial basis. Alternatively, provision may follow a more hybrid or mixed model. For purpose of this analysis, we do not make a distinction as to which mode of DDC production adopted. One might assume, however, that the more outsourced version given it is likely to rely on contractual arrangements that are more susceptible to monitoring and scrutiny from external bodies than internal processes and criteria parties use when designing new resources or recruiting new staff and volunteers.

The first and most basic stage of DDC is that the campaign team must have *access to, and see value in the accumulation of large quantities of individualized digitized voter data*. While the type of data used can vary, it should be of sufficient volume and granularity to allow for individualized voter targeting (offline and online) at scale. Typically, this would include demographic data such as age and sex and socially acquired characteristics like educational attainment, as well more indirectly observed traits such as personality type, likes or interests.¹ Information about individuals’ offline and online behaviours such as their commercial transactions and subscriptions, social media likes and posts, and geo-location data would also be valuable for targeting purposes. To accumulate these data campaigns require access to a range of sources that can be categorized into four main types. The first of these are public records, covering administrative data such as electoral registration information and census data. In addition, commercial databases compiled by private companies such as Experian also offer a range of consumer and marketing data on individual voters for a fee. Parties themselves may also compile their “in-house” voter records based on the results of local or national party contacting and specially commissioned opinion polls. Fourth, the internet opens up a rich new source of individual data. While some of this is free, more targeted options are available at cost from social media platforms through purchase of access to custom or lookalike audiences. In terms of how these data are accessed and stored this can also vary with some campaigns building their “in house” databases from source (first party data), while others rely more heavily or even exclusively on external service providers or data brokers (second and third party data). Finally turning to the organizational structures and support required for this first stage of DDC, staff or teams with expertise in data management and infrastructure development, i.e. software

engineers, computer programmers and IT specialists, would be key to its implementation.

Once acquired or accessed the data then need to *subject to interrogation by appropriate methods of analysis*. While this analytics work can support a range of campaign activities such as volunteer management and fundraising, as Dommett, Barclay, and Gibson (2023) point out, electoral mobilization is usually the priority. Stage 2 of DDC thus typically involves analysis of data produced in stage 1 to create voter profiles and select target audiences. The methods used can range from basic descriptive statistics to more advanced computational modelling of voter characteristics to predict their responses to different mobilizing or persuasive stimuli – a technique known as propensity scoring. At its most sophisticated this work is driven by the findings from social scientific studies of voter mobilization and insights from behavioural psychology, e.g. “nudge theory”. Given the size of the datasets involved, the methods used are most likely automated and quantitative, although there would also be demand for human coding of qualitative data in terms of digital texts. Staff and volunteers with a computing, data and/or social science background would provide the organizational expertise required for this stage of DDC. Typically they would be referred to as the “analytics” team.

The *insights generated in stage 2 are then applied to the campaign’s external communicational channels and the design and distribution of personalized messages to individual voters at scale*. Key tasks associated with this stage would be the creation and placing of micro-targeted ads on social media news feeds or sidebars using automated matching facilities. Platforms such as Facebook and Google are most commonly used at this stage, but campaigns, notably in the U.S.A., are now increasingly exploiting “over the top” digital streaming services such as Netflix, Amazon Prime, Hulu that go “over the top” of terrestrial or cable TV channels to send individualized ads to subscribers based on their viewing habits. As a less costly option, campaigns can also leverage the results from their data analytics work to generate “in house” or organic campaign posts, tweets, videos, and text messages that are specifically designed to be shared and reposted. While this might be directed at their grassroots it might also be targeted toward the new set of online “influencers” who occupy an important position of informal authority over the extensive networks of followers (Riedl et al. 2021). A core feature of this third stage of DDC is that it is evaluative and iterative. The results of the targeting are fed back into the types of data collected, analytical models and construction and delivery of the message to improve their accuracy and impact. At its most intensive, this evaluation uses experimental methods to systematically test the value of their appeals either after the fact, or ideally in real time. The organizational expertise associated with this stage of DDC would be located typically in the “digital creative” and advertising teams. Those with knowledge of the process of targeting itself, and particularly across the key platforms such as Facebook would also be vital at this stage.

The fourth and ultimate stage of DDC is that *data drives the decision-making*. The outcomes of the preceding data access and acquisition, analysis and communication stages are effectively “hard-wired” into the party structures and determine outcomes to on all major campaign decisions over and above any human judgements. In its extreme form, the “leadership” have no autonomy in allocating how and where and how the campaign targets its appeals and core resources. The key organizational staff here would be AI specialists that design the algorithms that underpin the new decision-making process

such that it uses prior information to learn, correct, and improve its outcomes. In the interim, one might see a new role for a “data interpreter” to emerge. This would be someone who has an understanding of how the AI machinery works, in terms of relating inputs to decision outputs and can explain this to senior organizational staffers. The process of automation could extend to shaping the priorities and decisions of the local parties and volunteers, as they are increasingly networked into central systems with mobile apps able to continuously upload data that monitors their activities and performance in meeting centrally set targets.

Drawing the stages together, we summarize the core traits and particular manifestations or indicators of each in tabular form. As well as simplifying and condensing the prior discussion, the table also helps to highlight the extent to which our operational definition provides the basis for a comparative and over-time understanding of DDC from a regulatory perspective but also for future academic analysis. In differentiating the generic “process” of DDC in stages from the particular activities that exemplify it we allow for the adaptation of the indicators to both cross-country variation and changes over time. While the former remain largely “fixed”, the latter are more flexible and can be updated.

Based on the prior discussion, we can also join these stages together to provide a new operationally focused definition of DDC as follows:

Data driven campaigning occurs when organizations have access to, and see value in the accumulation of large quantities of individualized digitized voter data, use computational and/or social scientific methods to analyse those data and apply the results of that process in an automated or deterministic way, to create and distribute individualized voter communication at scale.

This definition builds on and advances current definitions of DDC in that it identifies and “sorts” the activities and tools that are frequently cited as part of this new mode of campaigning into a more dynamic and organizationally-centred or bounded framework. In so doing, it also provides more specificity as to the core data, methods and goals that distinguish DDC from previous modes of campaigning.

Mapping regulatory attention to data-driven campaigning

The first part of this paper has drawn on the current literature to set out a new process-based, organizationally centred definition of DDC. The definition as [Table 1](#) shows allows for the allocation of a range of empirical indicators to each stage. These can be used to map the emergence of this new mode of campaigning across parties and party systems, and more specifically for our purposes, the attention it receives from governing bodies. The second section of the paper uses the indicators presented in [Table 1](#) as a means to map the regulatory environment that exists and is emerging around DDC in a specific governmental context. Specifically we show which components of DDC are gaining most legislative attention from European Union (EU) lawmakers, and conversely which are not under scrutiny. As noted above, the indicators as currently specified are not designed to yield a numerical score measuring extent of regulatory coverage of DDC. They do, however, form the basis for a more generating a quantitative ranking approach to governmental activity in regard to DDC. In a final section of the paper, we assess the implications of these “hot” and “cold” spots for future regulatory rollout within the EU and national policy context.

Table 1. Empirical Indicators of data-driven campaigning.

DDC stage	Indicators (e.g. in parentheses)
(1) Data Access	
1.a.i Data type	Geographic information (postcode, home address) Online contact information (email address) Mobile phone numbers Landline phone numbers Demographic information (sex, age, education) Social and recreational activities and interests Political outlook (likely to vote, party preference) Political activities (turnout, donation) Personality profile (big 5 typing)
1.a.ii Data source	Online behaviours (browsing, following, purchasing) Public records (electoral roll, census) Commercial vendors (Experian) Party / private (canvassing reports, internal opinion polls, focus groups) Social media platforms (Twitter) Software engineers; IT specialists
1.b Org expertise	
Stage (2) Data Analysis	
2.a Methods of analysis	Descriptive statistics (frequencies, cross-tables) Explanatory statistics (multivariate regression) Predictive statistics (forecasting district outcomes, vote propensity scores) Algorithm design, Machine learning, natural language processing
2.b.Org expertise	Data scientists; analytics team
Stage (3) Voter communication	
3.a. Message construction and distribution	Design and distribution of targeted ads on social media platforms (Meta) Design and distribution of targeted ads on streaming services (Netflix) Production and promotion of indirectly targeted shareable content (tweets) Experiments to evaluate and adjust online messages (A/B testing of ads) Digital “creatives”; outreach teams
3.b.Org expertise	
Stage (4) Decision-making	
4. a Integration of stages 1-3 into decision-making structures	Staff and volunteer activities tracked and managed by automated system Mandatory use of field and mobile apps Chief Information Officer (CIO) is campaign manager Final decisions rely on artificial intelligence outcomes of stages 1–3
4.b. Org expertise	Computing and AI specialists – Data interpreter

The selection of the EU as a case study of an emergent regulatory regime for DDC is made on both substantive and practical grounds. Given it is not possible to audit our DDC indicators against the full gamut of global legislation, we have selected a context that arguably offers an ideal reference point for this exploratory mapping exercise. First, the EU is widely regarded as a global leader in setting legal limitations on organizations use of citizens’ data (Newman 2018; Schwartz 2019). The triad of 2018 GDPR along with the more recently proposed DSA and DSA together form one of the most comprehensive and wide-ranging regulatory regimes developed to control the production and distribution of digital content in general.² Second, the recent announcement to extend the DSA to the regulation of online political advertising sets a new international standard that as we outline below, has direct applicability in relation to several core DDC practices. Finally, given the vast number of organizations and citizens subject to its remit, an EU system of regulation will be at the maximum end of the scale in terms of impact.

Current controls on DDC: identifying the “hot” spots

Below we review and explain how existing and forthcoming rules within the EU apply across the four stages of DDC. As become evident the main targets or “hotspots” of

attention centre on activities linked with stages 1, 2 and 3. We summarize this information through a summary version of Table 1 that includes the allocation of relevant legislation to each of the phases.

The main component of EU law relevant to regulating DDC is the 2018 General Data Protection Regulation or GDPR. While the passage of GDPR took place relatively recently, the EU's commitment to its citizens' rights to privacy and data protection is long-standing.³ In its current form, the GDPR has relevance for several aspects of DDC, most obviously with regard to stage 1 (1.a.i), i.e. access to voters' personal data. The central aim of the GDPR is to standardize and enhance protections on the use of EU citizens' personal data, defined in Article 4 as information relating to an identifiable living person. Most if not all of the items included in stage 1.a.i. are thus covered by GDPR, which means that organizations must obtain individuals' consent for the collection and processing of these data. Several items are given greater protection under Article 9 as "sensitive data", which means they are explicitly prohibited for data processing purposes. This includes individuals' racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, as well as biometric and genetic data.

Although these provisions might seem to place a heavy, and even insurmountable barrier on an organization's ability to conduct DDC (particularly the ban on processing of political opinions), closer inspection of the legislation highlights a number of "grey areas" in regard to the applicability of these restrictions, particularly in the electoral context. Notably, clauses I and (f) of Article 6 allow for non-sensitive personal data to be processed without a subject's consent, if there is a *public or legitimate interest* on the part of the data controller or a third party in doing so. For sensitive data, exceptions to the prohibition apply, if informed and explicit consent is given by the data subject to its use for a specified purpose, or where an organization has "regular" contact with an individual, or, again if it can be shown there is a "substantial public interest" in the processing of such data (Article 9, clauses a, d and g). While the extent to which European parties are exploiting these loopholes to date is unclear, reports by national government and privacy watchdogs have shown they are engaged in collecting a large quantity of protected and sensitive data on voters.⁴

One type of personal data listed in stage 1 that the EU have taken much more unambiguous action to restrict access to, are newer forms of digital data. The e-Privacy regulation (ePR) set to be enacted in 2023 will place much stricter controls on organizations' ability to collect information about individuals' online activities. The new rules are *lex specialis* to GDPR and as such supersede and codify the current "standard" protections afforded to such data. These have typically focused on ensuring permissions for the insertion of electronic identifiers such as cookies on users' browsers, which allow website owners to track their activity. The ePR dramatically expands these protections to include all information from individuals' browsers and any of their devices that connect to internet. Furthermore, and as a sign that lawmakers are seeking to tighten up the rules on consent, the new e-privacy legislation removes any reference to this being waived if there is a legitimate "public interest" in doing so.

A number of stages 2 and 3 DDC activities also come under the rubric of GDPR. Specifically the profiling and the micro-targeting of voters (items 2a. iv, vi and viii, 3a.i) are limited under Article 22 which states that individuals "have the right not to be subject" to decisions based these techniques, if they have applied in a wholly

automated manner, and have a “legal or similarly significant effect” on the data subject, i.e. that they reduce their capacity to exercise their legal rights. In the campaign context, an example of this might be a party’s use of algorithmic methods to target opponent supporters’ with false information about their preferred party to deter them from turning out to vote. Such an action would arguably place a restriction on an individuals’ free exercise of their right to vote. Article 22 also places further restrictions on stage 1 data collection, given that profiling techniques are often the basis for inferring non-directly observable traits such as personality type or interests.

In addition to GDPR, a tranche of more recent legislation has increased the EU’s regulatory oversight of DDC activities, particularly those associated with stages 2 and 3 concerning the profiling and placing of targeted ads on social media feeds using automated matching tools. The newly agreed Digital Services and Digital Markets Acts (DSA and DMA) grant significant new powers to the Commission to stem citizens’ exposure to illegal and harmful content online and open up competition in the provision of goods and services beyond the dominant social media companies. The DSA has the most direct relevance for DDC in that it imposes tough new requirements on producers and distributors of digital content to make their data processing and targeting techniques more transparent. In late 2021, these requirements were extended specifically to cover political advertising as part of the “Democracy and Integrity of the European Elections” legislative package. Under the proposed “Regulation on the Transparency and Targeting of Political Advertising”,⁵ parties and other campaign organizations must make clear when citizens are exposed to paid political content. Political content being defined influence the outcome of an election or referendum, a legislative or regulatory process or voting behaviour. If it is sponsored it must be clear who paid for it, how much was spent and why they are showing it, i.e. some explanation of the link between the advertisement and the relevant election or referenda must be provided. If the advert does not adhere to the new rules, it cannot be published (Chapter II, article 6). Such proposals thus have significant implications for stage 3 activities in terms of the design and distribution of paid ads by political parties in advance of an election.

As well as enhancing the transparency of the source, financing and purpose of the ad, the proposals also impose new requirements on providers to expose the targeting processes, i.e. stage 2 of DDC, underlying it. Specifically, sponsors must make clear that an ad is targeted, the criteria that have been used for that targeting, and whether any “amplification techniques” were used to widen its reach. Any internal policies in place regarding the use of such techniques must be made public (Chapter III, article 12). In a push to seal the potential loopholes that existed under GDPR, amendments were added by the European Parliament (EP) to the legislation that increased the restrictions on the use of personal data for profiling purposes, by explicitly ruling out the use of sensitive data in targeted advertising, and removing any exception to in relation to the public interest. In addition, the use of profiling techniques in relation to minors was expressly prohibited.

Although it appears that micro-targeting on social media platforms is currently under most scrutiny from EU regulators, there are signs that other communication services may be moving into the frame. Specifically, ePR legislation and the 2018 European Electronic Communications Code have both been expanded to include reference to “over-the-top” or OTT services, which, as the name indicates, operate over the top of the existing network offered by a service provider. Thus far, the EU’s attention has centred on services

that allow for inter-personal communication such as Skype, WhatsApp and Zoom and function through voice over protocols. As such, the primary focus of regulation has been to protect the privacy of users' communication and promote competition among suppliers of such services within Single Market. Extension to include video services such as Netflix and Hulu and impose DSA related transparency requirements for political advertising, however, would form an obvious next step, should these channels begin to carry micro-targeted ads.

In terms of the current controls on DDC, therefore, as our review has revealed, almost all stages of DDC, with the exception of stage 4 are covered in some way by current or pending legislation. To summarize the main areas of DDC that are subject to current or pending regulation from EU and highlight those that are not under scrutiny we map the relevant legislative initiatives to the various substages in [Table 2](#).

While stage 1, and particularly access to data has been under the regulatory radar for some time, stages 2 and 3 form particular foci of government attention, with political organizations facing increasing restrictions on their voter profiling activities and use of paid advertising on social media channels. The new regulations on transparency and targeting of political advertising, following "trilogue" between the European Commission, Council and Parliamentary approval, are due to take effect in all member states from April 2023. National governments then have 1 year to set out how they will be implemented. Specifically, they are required to allocate responsibility for monitoring use of the profiling techniques to their data protection authorities, designate competent national authorities to monitor compliance with their other obligations and specify the sanctions to be imposed for violations.

What is left out? Identifying the "cold spots"

The preceding review has shown that a substantial number of DDC relevant activities are, or soon will be, subject to regulation by a major international governance body. It also reveals there are a number of gaps in the extent of EU oversight. Some of these exclusions can be explained by the presence of existing national regulations and the need to allow for flexibility at the "local" level. Access to some of the main data sources relevant for the first

Table 2. EU legislation relevant to data-driven campaigning by stage.

DDC stage	Component	Relevant legislation
1.a.i Data type	Geographic information Online contact information Mobile phone numbers Landline phone numbers	GDPR Article 4, personally identifiable data
1.a.i Data type	Demographic information Political outlook Political activities	GDPR Article 9 "sensitive" data
1.a.i. Data type	Personality /psychological profile	GDPR Article 22
1.a.i Data type	Online activities	e-Privacy regulation (ePR) lex specialis to GDPR; European Electronic Communications Code (EECC)
2.a Methods of analysis	Propensity scoring	GDPR, Article 22; Digital Services Act (DSA), particularly "Regulation on the Transparency and Targeting of Political Advertising" Ch.III, Article 12
3.a Message construction and distribution	Social media /streaming services ads	DSA, particularly "Regulation on the Transparency and Targeting of Political Advertising", Ch II, Article 6; EECC; ePR

stage of DDC for example is largely determined by national governments. Specifically the rules on use of public and administrative data such as the electoral roll differ among EU countries. In addition, the quality and usability of the data sources vary according to national requirements regarding if voter registration is compulsory or voluntary, and how records are stored i.e. in digitized, biometric or paper-based format. Beyond the official sources of DDC data, however, it does appear that the EU are looking to standardize and limit organizations access to newer sources of personal information such as that generated via the internet and social media platforms. The new ePR legislation and updates to the European Electronic Communications Code (EECC) in particular will impose significant restrictions on campaigns and other groups' ability to access individual data from a range of new electronic communication services including WhatsApp, Facebook Messenger and Skype. Essentially the goal is to ensure any interactions or communication engaged in on these platforms are subject to the same level of confidentiality of communications as traditional telecom operators.

Other gaps in current EU provisions toward DDC appear to be more obvious "cold spots" i.e. areas that regulators have for now, chosen not to focus on. We can identify two main areas of omission. The first being the internal component of parties' self-regulation in relation to DDC. Here, we can see two main gaps. The first is in regard to management of the new skills and expertise that will flow into political parties as they expand their DDC efforts, i.e. items 1.b., 2.b. and 3 b. as well as stage 4 more generally. While such individuals will no doubt be well trained in the data science of micro-targeting and use of algorithms, the extent to which they are required to be aware of the legal and ethical limitations on the use of these techniques is currently not a matter for public concern, as viewed under the EU's existing regulatory framework. Again, this gap may be based on an understanding that the governance of intra-organizational practices is best addressed at the national level. A second gap in oversight at the intra-organizational level concerns the monitoring and tracking of volunteers' performance and activity. Although one might consider such actions to be covered by the same rules that apply to parties' use of voters' personal data, there would appear to be waived for members and registered supporters under current GDPR rules since they have ongoing "regular" contact with the party. As such, they can be seen to have consented to their data being processed, at least for the purpose of campaign coordination and planning.

A second major area of DDC activity that appears (for now) to have escaped EU regulation relates to stage 3 and the production of non-sponsored or organic content and indirect voter targeting (item 3a.ii). The exclusion of such content from the new transparency requirements for political advertising in effect allows third parties to disseminate election material on behalf of a candidate or party targeted toward certain voters. This could involve a coordinated effort to release campaign content through non-official sources such as popular bloggers and political influencers or bots. Future efforts to bring unpaid content under the DSA regulations, however, cannot be ruled out and look increasingly likely. The question of how far the current rules can be extended to cover non-remunerated content is one that has prompted increasing debate among legal experts at the EU⁶ and national level.⁷ Also, the fact that the current legislation has already been extended to include "issue-based ads" is seen as broadening its remit to cover a much wider range of content produced by citizens and pressure groups, outside of election periods.

The regulation of data-driven campaigning: looking ahead

We conclude by drawing on the findings of our heat mapping exercise to anticipate how the regulation of DDC is likely to develop in the future, focusing primarily on the case of the EU. Given the increasing number and range of activities that have moved under its policy lens, we would expect continuing attention to existing hot spots and further expansion into currently “cooler” areas, such as the intra-organizational sphere to take place. In terms of how this might occur, we outline two potential pathways – a “kaleidoscopic” and “designed” approach. The former follows a more gradualist and legacy-driven logic whereby a wider range of DDC activities increasingly fall under the scope of existing legislation. The latter constitutes a more interventionist approach in which new institutions and bespoke regulations are designed to control and limit its use. To develop these trajectories, we draw loosely on theories of internet governance and particularly earlier work by Solum (2008) that identified several models for how cyberspace could be “ruled”. While the majority of these models were actually “free” from state control⁸, two were rooted to varying degrees, in existing governing structures. The more radical “transnational” version imposed controls over the internet through the design and coordination of new specialist set of supra-national institutions. In contrast the “national government” model worked more simply by extending a country’s existing legislation to encompass a range of new internet-based activities, e.g. publishing contracts and defamation laws. Leaving aside the territorial assumptions of each model, and accepting that governance in this sphere cannot be directly equated with regulation (Hofmann, Katzenbach, and Gollatz 2017), the contrast offered by Solum between an interventionist approach to designing internet protocols, and the gradual extension of legacy systems provides a very useful basis for grounding and developing our two models of DDC regulation. Furthermore, given recent claims for a “return of the state” by internet governance scholars (Haggart, Scholte, and Tusikov 2021) and the adoption of a more interventionist stance by leading national politicians including the French President,⁹ renewing debate on differing approaches to governing digital realm appears to be particularly relevant to current debates.

Turning first to the “kaleidoscopic” model, we argue this constitutes the current default model in the EU and thus presents the most likely pathway for future regulatory activity. Here, DDC is increasingly controlled by a combination of EU and national law in a piecemeal and incremental fashion as more activities fall under the ambit of existing legal provisions and norms. This process of absorption is perhaps most evident at the EU level in the regulation of stage 1 DDC resources. As campaigns have amassed and processed increasing amounts of voters’ personal data, their actions are increasingly subject to scrutiny by the EU under the provisions of GDPR. GDPR as discussed itself forms the legal codification of the deeper founding commitment of the EU to protect individuals’ privacy and personal information as set out in the 1950 European Convention of Human Rights. This incorporation can also arguably be seen in the controls now being applied to stage 3 activities. The regulation of the content, targeting and transparency of online political advertising has emerged from the EU’s ongoing broader legislative commitment and programme to combat the spread of misinformation and curb the powers of big tech. These initiatives can and have been traced back by proponents of the new rules to the founding principles of the EU Charter that its citizens will receive

information free from “interference”.¹⁰ According to the kaleidoscopic logic, therefore, the regulation of DDC occurs mainly as a “by-product” of existing EU and relevant national laws, and will continue to do so as more component activities successively fall under existing areas of regulatory focus.

The “designed” mode, as one might expect, takes a more systematic, interventionist and potentially transnational approach to the task of DDC regulation, and sets out a new comprehensive “purpose-built” system of control to directly limit and even prohibit its full expression. Essentially, governing elites or those within the relevant executive agencies increasingly perceive the problems presented by the un- or under-regulated use of DDC and seize the opportunity this presents for review and reform of campaign and election regulation, with the view toward establishing a new and more centralized regime. While such a regime would obviously be context sensitive, one could envisage here the establishment of a new powerful cross-sector agency that assumes the regulatory and oversight functions previously exercised by other bodies such as the Electoral Commission, the Data Protection agencies, and the Media and Advertising standards regulator. In the most radical version, these regulatory bodies could be restructured and merged into new Electoral Campaign and Data Commission.

As noted, the kaleidoscopic model appears, at least *prima facie*, to describe the current approach within the EU toward DDC regulation. Despite its supranational status, the EU remains ultimately a union of sovereign states and any new rulings are subject to national governments’ consent, and alignment with a heterogeneous body of national laws. The new regulations on transparency and targeting of political advertising are devolved to national authorities, notably the data protection agencies who will set and administer the new system of financial penalties. That said, beyond these structural constraints, and returning to our introductory discussion about the current alarm felt about the spread of these techniques, clearly contextual factors that may intervene and enhance the appeal of the designed mode at a supra-national level. Specifically, an escalation of public and particularly elite perceptions of the potential harms associated with DDC and the extent to which it seen as a subversive force in politics will likely fuel demands for more rigorous reforms of the current system. External events and particularly new crises over data breaches and voter profiling scandals during elections are likely to trigger demands for greater government intervention. The exposure of anti-democratic practices and attempts to manipulate outcomes emerges in relation to the EP elections in 2024 could help forge momentum for a more “joined up” solution in the form of a new EU Election and Data Protection agency. As well as monitoring voter profiling and targeting, this body could also impose new procedures for monitoring and regulation of parties’ internal practices in this area, e.g. the mandatory auditing of European parties data processing and formal accreditation and training of staff to ensure appropriate standards are followed. By contrast, in the absence of further scandals, it is more likely that a more “normalized” or even sceptical interpretation of DDC would be maintained. The incorporation of new more sophisticated processing and profiling techniques would form part of an inevitable progression and sophistication of “standard” campaign practices, and reports of voter surveillance and manipulation are dismissed as over-stated. In such a context, “legacy” frameworks will be considered as fit-for-purpose and extended for regulatory purposes.

Conclusion

While we have focused here on a specific task of mapping the scope and structure of current and future regulation of DDC to the case of the EU, we conclude by drawing out the relevance of our analysis and utility of our operational definition for further studies. First, it is clear that the indicators we identify can be applied to the national regulatory context. While this would form a useful exercise in its right, it would also allow us to pinpoint where and how national legislation is lacking, or conversely extends controls on DDC beyond international law. A brief survey of activity among European countries shows a number of new initiatives being proposed or enacted to protect the citizenry against online harms of misinformation and manipulation.¹¹ The extent to which these will provide a basis for strengthening powers to control political campaigning, and particularly messaging content is an important question for future study.

Second, in mapping these new regulatory interventions we highlight an important new interface for the study and practice of campaigns. Parties' embrace of DDC, as we have shown, moves them increasingly under the radar of existing and new regulations designed to protect citizens' privacy and provide more transparency in the content they are exposed to online. While scholars have always been cognizant of the importance of the financial and broadcasting rules governing elections in explaining parties' campaign activity, there is now an increasing need to extend that understanding to incorporate privacy and personal data protection law. Such provisions are likely to have significant implications for the extent and nature of parties' communication both within and between elections.

Finally, beyond the regulatory context, we envisage this exercise to operationalize the concept of DDC and produce observable indicators of its practice will advance more quantitative analysis of its adoption in the national and international context. [Table 1](#) provides a formative basis for the construction of numerical index on which not only legislative frameworks can be evaluated for their coverage of DD, but through which individual parties and candidates can be scored for their coverage or uptake. Furthermore in separating what are considered as the generic and fixed elements of DDC from its more context dependent and flexible components, we also open up the potential for more inter-party and cross-national analysis to be conducted. Measuring which organizations and countries are "ahead" or behind in terms of DDC adoption will open up new opportunities to identify both the drivers behind this new mode of electioneering and its electoral and wider democratic consequences.

Notes

1. For further insight into the extensive range of data typically contained in voter files held by U.S. Campaigns, see the updated models for 2021 published by L2 and Haystaq <https://haystaqdna.com/wp-content/uploads/2021/09/L2-National-Models-User-Guide-2021.pdf>
2. Jillian Deutsch E.U. Lawmakers Propose New Rules to Control Online Content <https://time.com/6128245/eu-online-content-regulations-internet/> *Time Magazine* December 13th, 2021.

3. GDPR updates and extends the 1995 Data Protection Directive. Prior to this, however, these rights were enshrined in in the 1950 European Convention of Human Rights (Article 8) and recommitted to and extended in the 2000 European Charter of Fundamental Rights (Article 7) which sets out the explicit right to the protection of personal data (Article 8).
4. In the UK, the Information Commissioners Office (ICO) investigated parties compliance with GDPR following the 2017 election and issued 11 formal warning letters highlighting the risks their methods posed to voter privacy <https://ico.org.uk/media/action-weve-taken/2260271/investigation-into-the-use-of-data-analytics-in-political-campaigns-final-20181105.pdf> A Open Rights Group report published in 2020 found evidence of extensive use of personal data by UK parties and a worrying lack of understanding among parties about their obligations and rights in this regard <https://www.openrightsgroup.org/app/uploads/2020/07/200619%E2%80%94org%E2%80%94report.pdf>
Elsewhere, investigation of French campaigns has also indicated that French parties have exploited loopholes in GDPR to collect and use voters personal data without their consent. <https://ourdataourselves.tacticaltech.org/posts/overview-france/>
5. Regulation on the Transparency and Targeting of Political Advertising (COM (2021) 731 final). https://ec.europa.eu/info/sites/default/files/3_1_178086_pol-ads-annex_en.pdf
6. Expert analysis by Prof Lorna Woods of the proposed regulations <http://eulawanalysis.blogspot.com/2021/11/the-proposed-eu-regulation-of-political.html> 21.11.21 noted that “although ‘adverts’ are the specified target of the new rules there is some ambiguity as to whether only financially remunerated ads are in scope or they extend to non-remunerated content. While the definition of a political advertising service might imply service for remuneration (see recital 29), the definition of the type of content contained in political advertising itself does not include any element of remuneration”.
7. In a formal consultation exercise, the UK Electoral Commission invited comment on whether organic election content should be subject to the same requirements for transparency as paid advertising. <https://www.electoralcommission.org.uk/who-we-are-and-what-we-do/changing-electoral-law/transparent-digital-campaigning/report-digital-campaigning-increasing-transparency-voters>
8. These being a reliance on pure self-governance, market forces, or the technical infrastructure and code that underpinned the internet as a whole.
9. Speech to the Internet Governance Forum www.intgovforum.org/multilingual/content/igf-2018-speech-by-french-president-emmanuel-macron. Accessed 4th July 2022.
10. Article 11 of the European Charter of Fundamental Rights. *Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers.*
11. Recent proposals set forward by the Belgian CSA https://www.epra.org/news_items/belgium-the-csa-calls-for-new-legal-duties-for-social-networks-and-online-platforms-to-fight-online-hate-speech explicitly point to the fact they emulate the NetzDG law in Germany and the draft law on online hate in France (Proposition de loi “Avia” visant à lutter contre la haine sur internet) currently under debate in the French Parliament. In the meantime, in the UK, following the public consultation on the Online Harms White Paper, the Government has announced that Ofcom will be in charge of regulating the internet, with potential criminal penalties for executives of internet firms failing to protect users from “harmful and illegal content” online.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by European Research Council [Grant Number 833177].

Notes on contributors

Rachel Gibson is a professor of politics at the University of Manchester. She currently leads the ERC Advanced Investigator grant project Digital Campaigning and Electoral Democracy (DiCED) and is a principal investigator of the Norface project examining The Intended and Unintended Consequences of Data-Driven Campaigning. She has published widely on the topic of digital campaigning including her 2020 book, *When the Nerds Go Marching In* with Oxford University Press.

Esmeralda Bon is the lead research associate for the ERC Advanced Investigator grant project Digital Campaigning and Electoral Democracy (DiCED) at the University of Manchester. Her thesis and published work has focused on the drivers and impact of established and new political elites' online political communication. Her current research interests centre on the growing role of political influencers in elections and developing methodologies to study this.

Andrea Römmele is Dean of Executive Education and Professor of Communication in Politics and Civil Society at the Hertie School. Her research interests are comparative political communications, political parties and public affairs and she has published widely on these topics in both peer review journals and monographs. She appears regularly as a political commentator on news and current affairs programmes.

References

- Adcock, R., and D. Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95 (3): 529–546. <https://doi.org/10.1017/S0003055401003100>
- Anstead, N. 2017. "Data-driven Campaigning in the 2015 United Kingdom General Election." *The International Journal of Press/Politics* 22 (3): 294–313. <https://doi.org/10.1177/1940161217706163>
- Baldwin-Philippi, J. 2017. "The Myths of Data-Driven Campaigning." *Political Communication* 34 (4): 627–633. <https://doi.org/10.1080/10584609.2017.1372999>
- Bennett, C. 2016. Is Your Neighbor and Democrat or a Republican? Lateral Voter Surveillance and the Political Culture of Modern Election Campaigns. Available at SSRN: <https://ssrn.com/abstract=2776308> or <http://doi.org/10.2139/ssrn.2776308>.
- Blumler, J. G. 2016. "The Fourth Age of Political Communication." *Politiques de Communication* 1 (1): 19–30. <https://doi.org/10.3917/pdc.006.0019>
- Borgesius, F. J. Z., J. Möller, S. Kruike-meier, RÓ Fathaigh, K. Irion, T. Dobber, B. Bodo, and C. De Vreese. 2018. "Online Political Microtargeting: Promises and Threats for Democracy." *Utrecht Law Review* 14 (1): 82–96. <https://doi.org/10.18352/ulr.420>
- Dobber, T., Ó Fathaigh, and F. Zuiderveen Borgesius. 2019. "The Regulation of Online Political Micro-Targeting in Europe." *Internet Policy Review* 8 (4). <https://doi.org/10.14763/2019.4.1440>
- Dommett, K., A. Barclay, and R. Gibson. 2023. "'Just What Is Data-Driven Campaigning?: A Systematic Review'." *Information Communication and Society*, 1–22. <https://doi.org/10.1080/1369118X.2023.2166794>
- Endres, K., and K. J. Kelly. 2018. "Does Microtargeting Matter? Campaign Contact Strategies and Young Voters." *Journal of Elections, Public Opinion and Parties* 28 (1): 1–18. <https://doi.org/10.1080/17457289.2017.1378222>
- Esser, F., and B. Pfetsch. 2020. "Comparing Political Communication: A 2020 Update." *Comparative Politics*, 336–358. <https://doi.org/10.1093/hepl/9780198820604.003.0019>
- Gorton, W. A. 2016. "Manipulating Citizens: How Political Campaigns' Use of Behavioral Social Science Harms Democracy." *New Political Science* 38 (1): 61–80. <https://doi.org/10.1080/07393148.2015.1125119>
- Haggart, B., J. A. Scholte, and N. Tusikov. 2021. "Introduction: Return of the State?" In *Power and Authority in Internet Governance*, edited by B. Haggart, J. A. Scholte, and N. Tusikov, 1–12. Abingdon, Oxon: Routledge.

- Harker, M. 2020. "Political Advertising Revisited: Digital Campaigning and Protecting Democratic Discourse." *Legal Studies* 40 (1): 151–171. <https://doi.org/10.1017/lst.2019.24>
- Hofmann, J., C. Katzenbach, and K. Gollatz. 2017. "Between Coordination and Regulation: Finding the Governance in Internet Governance." *New Media & Society* 19 (9): 1406–1423. <https://doi.org/10.1177/1461444816639975>
- Kefford, G., K. Dommett, J. Baldwin-Philippi, S. Bannerman, T. Dobber, S. Kruschinski, S. Kruikeimeier, and E. Rzepecki. 2022. "Data-driven Campaigning and Democratic Disruption: Evidence from Six Advanced Democracies." *Party Politics*, <https://doi.org/10.1177/1354068822108403>.
- Kreiss, D. 2016. *Prototype Politics: Technology-Intensive Campaigning and the Data of Democracy*. New York: Oxford University Press.
- Kruikeimeier, S., M. Sezgin, and S. C. Boerman. 2016. "Political Microtargeting: Relationship Between Personalized Advertising on Facebook and Voters' Responses." *Cyberpsychology, Behavior, and Social Networking* 19 (6): 367–372. <https://doi.org/10.1089/cyber.2015.0652>
- Kruschinski, S., and A. Haller. 2017. "Restrictions on Data-Driven Political Micro-Targeting in Germany." *Internet Policy Review* 6 (4): 1–23. <https://doi.org/10.14763/2017.4.780>
- Lewandowsky, S., L. Smillie, D. Garcia, R. Hertwig, J. Weatherall, S. Egidy, R. E. Robertson, et al. 2020. Technology and Democracy: Understanding the Influence of Online Technologies on Political Behaviour and Decision-making. Available at https://pure.mpg.de/rest/items/item_3277241/component/file_3277242/content Accessed 08.08.21.
- McEvoy, M. 2019. "Big Data and Democracy: A Regulator's Perspective." *Internet Policy Review* 8 (4).
- Munroe, K. B., and H. D. Munroe. 2018. "Constituency Campaigning in the Age of Data." *Canadian Journal of Political Science/Revue Canadienne de Science Politique* 51 (1): 135–154. <https://doi.org/10.1017/S0008423917001135>
- Newman, A. L. 2018. *Protectors of Privacy*. Ithaca, NY: Cornell University Press.
- Nickerson, D. W., and T. Rogers. 2014. "Political Campaigns and big Data." *Journal of Economic Perspectives* 28 (2): 51–74. <https://doi.org/10.1257/jep.28.2.51>
- Persily, N. 2017. "The 2016 US Election: Can Democracy Survive the Internet?" *Journal of Democracy* 28 (2): 63–76. <https://doi.org/10.1353/jod.2017.0025>
- Riedl, M., C. Schwemmer, S. Ziewiecki, and L. M. Ross. 2021. "The Rise of Political Influencers— Perspectives on a Trend Towards Meaningful Content." *Frontiers in Communication* 6), <https://doi.org/10.3389/fcomm.2021.752656>.
- Römmele, A., and R. Gibson. 2020. "Scientific and Subversive: The Two Faces of the Fourth Era of Political Campaigning." *New Media & Society* 22 (4): 595–610. <https://doi.org/10.1177/1461444819893979>
- Schwartz, P. M. 2019. "Global Data Privacy: The EU way." *NYUL Rev.* 94: 771.
- Solum, L. B. Models of Internet Governance (2008). Illinois Public Law Research Paper No. 07-25, U Illinois Law & Economics Research Paper No. LE08-027, Available at SSRN: <https://ssrn.com/abstract=1136825>.
- Stromer-Galley, J. 2014. *Presidential Campaigning in the Internet Age*. New York: Oxford University Press (1st Edition).
- Turow, J., J. King, C. J. Hoofnagle, A. Bleakley, and M. Hennessy. 2012. Americans Reject Tailored Advertising and Three Activities That Enable It. Available at SSRN 1478214.
- Witzleb, N., M. Paterson, and J. Richardson, eds. 2019. *Big Data, Political Campaigning and the Law: Democracy and Privacy in the Age of Micro-Targeting*. Abingdon, Oxon: Routledge.