Media and Communication
2024 • Volume 12 • Article 8735
https://doi.org/10.17645/mac.8735

ARTICLE

Open Access Journal

Data-Driven Maintaining: The Role of the Party and Data Maintenance in the US Context

Jessica Baldwin-Philippi

Department of Communication and Media Studies, Fordham University, USA

Correspondence: Jessica Baldwin-Philippi (jphilippi@fordham.edu)

Submitted: 23 May 2024 **Accepted:** 2 October 2024 **Published:** 31 October 2024

Issue: This article is part of the issue "Data-Driven Campaigning in a Comparative Context: Toward a 4th Era of Political Communication?" edited by Katharine Dommett (University of Sheffield), Rachel Gibson (University of Manchester), Sanne Kruikemeier (Wageningen University & Research), Sophie Lecheler (University of Vienna), Esmeralda Bon (University of Manchester), and Stephanie Luke (University of Sheffield), fully open access at https://doi.org/10.17645/mac.i457

Abstract

Political campaigning in the US is unique in the global context for its lack of attention to the role of the party, largely due to the centrality and power of campaigns. In the US context, successful data-driven campaigning (DDC) has often been covered by the press and analyzed by US scholars as an innovative campaign creating new tools and new tactics (and earning more media coverage for them). This research investigates the oft-ignored role of party organizations in DDC in the US, and in doing so, highlights the invisible work of data maintenance that is their purview. Methodologically, it brings together interviews with staffers from both party organizations and campaigns with thematic analysis of news coverage to answer questions about how the data-driven practices of parties versus campaigns differ, how parties' data work is (and is not) covered, and what, in staffers' views, contributes to such coverage. Ultimately, this research highlights how a lack of attention to party organizations' work has gone hand in hand with a lack of attention to maintenance work in both academic and public discussions of DDC.

Keywords

big data; campaigns; data-driven campaigning; maintenance; political campaigning; political parties

1. Introduction

In the decade since academic attention to data-driven campaigning (DDC) has taken off, the field has increasingly focused on a variety of international contexts (Dommett et al., 2021; Dommett, Kefford, & Kruschinski, 2024; Kefford et al., 2023; Roemmele & Gibson, 2020). Still, much of the foundational work in

the field centers on the US case (Baldwin-Philippi, 2019; Hersh, 2015; Karpf, 2016; Kreiss, 2016; Nickerson & Rogers, 2014; Stromer-Galley, 2014). Thus, this early work has often set a standard for comparative analysis (Gibson, 2020). One difficulty of engaging in such comparative work, however, is the mismatch in how US-focused scholarship has overwhelmingly focused on the role of candidate-centered campaigns, while studies of DDC in advanced democracies (indeed, all of the international examples above), center the role of the party. While it's relatively common knowledge that parties have databases that provide campaigns with lots of data, questions about which types of data practices US party organizations, as opposed to campaigns, engage in have garnered less attention from both academics and the press. In order to investigate what kinds of data work US party organizations do and how it gets covered, I bring together two methodological approaches. First, I conduct interviews with political professionals who specialize in data. I also engage in thematic analysis of news coverage of data campaigning. In both analyses, I focus attention on what types of data work—maintaining, modeling, targeting, and testing—party organizations engage in.

From the interviews, I find that political professionals describe the work done in national party organizations as focusing overwhelmingly on maintenance and modeling, and that maintenance, in particular, is seen as work that is unrecognized and unique to party organizations. From news coverage, I find that coverage of maintenance by party organizations is often overshadowed by coverage of campaigns and that maintenance is often described as a supportive and passive act of "providing" data, rather than actively cleaning, acquiring, standardizing, or creating new data. Overall, this work attempts to bring greater attention to the maintenance work done by party organizations in the wider campaign-focused US context.

2. Literature Review

2.1. Campaign-Centric Analysis of DDC in the US

For roughly the past century, US politics has fundamentally focused on the candidate-centered campaign (Arbour, 2014). While parties may hold power to surface and squelch nominees in primary elections (Cohen et al., 2008), even this stronger assertion of parties' power centers the subsequent candidate campaign as the main focus of electoral attention in general elections. So much so that the term "campaign" generally implies the separate candidate organization that is legally, financially, and organizationally separate from party organizations. This contrasts with the continued power of parties in other advanced democracies and systems of parliamentary rule, as:

> Parties in the United States certainly do provide a common label under which candidates run for office. However…the parties are not strong links between voters in elections and the officials they elect to represent them in party government, as the term is understood in other countries. (Dwyre, 2010, p. 28)

As such, instead of robust, centralized organizations, parties in the US are best understood loose networks or constellations of actors involving not only party committees (e.g., the Democratic National Committee [DNC] and Republican National Committee [RNC], Democratic Senatorial Campaign Committee [DSCC], or National Republican Senatorial Committee [NRSC], etc.) but also the attendant campaigns, private consulting firms, and even activist and advocacy organizations that make up the partisan ecosystem (Gerken & Fishkin, 2014; Masket & Noel, 2021). As a result of their relatively removed status, national party

organizations have developed to "play a supportive role, offering resources and services to candidates who seek their help" (Galvin, 2012).

This is particularly true in a digital context, as "the candidate-centric electoral process largely relegated parties to a supporting role as campaign organizations dictated media strategies" (Owen, 2013, p. 348). This "supportive" function, often maligned as powerlessness in political science literature, also deserves attention as what I will describe below as important "maintenance" work. In landmark US-based studies of DDC, scholars have focused on data infrastructure like databases used for targeting potential voters (Hersh, 2015; Kreiss, 2016; Nielsen, 2012), and positioned party organizations as one of the many actors in that space. As Kreiss's (2016) *Prototype Politics* centers the organizational development of data infrastructure, it focuses on the synergies between party organizations, campaigns, and third-party actors like for-profit tech consultancies to tell the story of how databases like the Data Center and VoteBuilder came to be. By focusing on synergies, it gives less attention to the ways the work done by parties may be fundamentally different than that done by campaigns. Hersh (2015) deeply investigates the types of data that make up the foundational data infrastructure of the left—including party-affiliated databases (NGP-VAN, which although a private company, makes the DNC's data useable by campaigns), as well as that of outside consultants (Catalist). Within the US context, outside of these two bodies of work, research on data campaigning focuses overwhelmingly on campaigns (Baldwin-Philippi, 2017, 2019; Kruikemeier et al., 2022; Nickerson & Rogers, 2010). As a result, accounts of DDC in the US remain incomplete due to the lack of focus on party organizations as separate entities. In doing so, they underplay the importance of parties in the US as key, if hidden, agents of delivering campaigns. Because analyses in other countries often center on the party, an increase in attention to the role of parties in the US also enables greater comparative analysis.

2.2. Goals of Data-Campaigning: Maintaining, Modeling, Targeting, and Testing

As the field of DDC has developed, many scholars have tried to assert "what counts" as DDC. One robust definition, created by conducting a descriptive meta-analysis of existing research on DDC, argues that it "relies on accessing and analyzing voter and/or campaign data to generate insights into the campaign's target audience(s) and/or to optimize campaign interventions. Data is used to inform decision-making in either a formative and/or evaluative capacity" (Dommett Barclay, & Gibson 2024, p. 2). Other definitions overlap but add additional focus on practices. Kefford (2021) describes DDC as "a set of inter-locking practices and processes which includes collecting data, building models of the electorate, creating supporter and persuadability scores, [and] segmenting and targeting voters at the individual level" (p. 6). Roemmele and Gibson (2020) argue that DDC involves developments in data infrastructure, a networked approach to voter communication, enhanced targeting, and internationalized campaigns. Baldwin-Philippi (2019) has argued that practices of targeting and testing are foundational to DDC. From the above definitions, four types of practices emerge as useful frameworks by which to assess what campaigns and party organizations are doing, and how that work is understood by the public and media: maintaining, modeling, targeting, and testing.

Maintaining is the work of cleaning, updating, and standardizing databases, assessing the use value of various datapoints, and collecting new data. Because campaigns rely on shared data that comes from a variety of data sources and can quickly become outdated, this work is integral to all political organizations' ability to engage in practices of modeling and is also directly related to efforts to target. This definition combines Kefford's (2021) emphasis on "collecting data," with Roemmele and Gibson's (2020) focus on infrastructure.

Beyond the realm of political communication, maintenance studies, a subfield within science and technology studies, has argued that although innovation and maintenance are intertwined: "Maintenance and repair, the building of infrastructures, the mundane labour that goes into sustaining functioning and efficient infrastructures, simply has more impact on people's daily lives than the vast majority of technological innovations" (Russell & Vinsel, 2016). Maintenance, they argue, has long gone ignored—by academics, the broader public, and the media that brings attention to such subjects–despite its centrality to all innovation. Scholarship on DDC and digital campaigning more broadly has long focused on innovative and cutting-edge practices, and continues to overlook maintenance risks ignoring the constantly ongoing work that is necessary to continue to develop new innovations. Dommett, Barclay, and Gibson's (2024) exhaustive definition above even leaves maintenance out, instead focusing on what campaigns do with already-formed databases.

Modeling focuses on the work of creating and refining new data points and algorithms that campaigns and parties can use to explain and predict the electorate in a variety of ways, from who is likely to vote versus who needs encouragement, or who is likely to be persuaded. This category of practice combines Kefford's (2021) two DDC qualities of building models of the electorate and creating supporter scores and Dommett, Barclay, and Gibson's (2024) focus on data segmentation, modeling, and behavioral monitoring. These models are often key to a wide variety of campaigns' and parties' strategy development, from when and where to deploy resources in canvassing labor to how to target ad buys.

Targeting has been defined by Baldwin-Philippi (2019, p. 2) as "using data to decide which messages go to what potential voters at what time during the campaign." This combines Kefford's (2021) focus on segmenting and targeting, Roemmele and Gibson (2020) focus on enhanced targeting, and Dommett, Barclay, and Gibson's (2024) definition of using data to develop target audiences.

Testing has also been defined by Baldwin-Philippi (2019, p. 2) as "empirically measur[ing] how well messages perform against one another and us[ing] that information to drive content production." This connects with Dommet et al's (2024) definition of formative assessment as well as "optimizing" campaign decisions, as A/B testing is often referred to as optimization testing.

In order to investigate what types of data work party organizations engage in, and how that compares to candidate-based campaigns, I take two different qualitative approaches. First, I focus on the perspectives of professionals themselves and investigate the ways they describe their work. Second, I turn to journalistic coverage of DDC to understand what work done by parties is actually explained to the public. As news coverage is a major way the public understands the intricacies of the actual work that goes into campaigning, there is value in combining professionals' own perspectives with the more generalist overviews of such work that news coverage is more likely to contain. Ultimately, I ask two research questions about these practices of data campaigning:

> RQ1: How do the focuses of political professionals who specialize in data campaigning in a national party setting differ from those who specialize in data campaigning within campaigns and consultancies, especially with regard to questions of data maintenance?
>
> RQ2: What does journalistic coverage of maintenance work done by parties look like?

Together, these approaches help reveal a holistic picture of data work that is both perceived as overlooked by party staffers and often is left out of news coverage.

3. Methods

To assess the political professionals' own accounts of their work, I conducted 15 interviews from 2019–2020, all with leaders in the field who were promised anonymity—people who ran data or digital teams in presidential elections, national party committees, or are founders or partners in firms that specialize in data campaigning from both the Democratic and Republican ecosystems. Descriptions of the professional experience of the interview subjects can be found in Appendix A of the Supplementary File. Of the 15 interview participants, nine came from the Democratic ecosystem and six from the Republican ecosystem. Five total staffers spent over three years doing data work in their respective national party organizations (three DNC, two RNC), six spent over a year (four DNC, two RNC), and four have never worked in their national party organization. Semi-structured interviews ranged from 30–66 minutes, with an average time of 43 minutes. These interviews were all conducted prior to the analysis of news coverage.

In order to investigate news coverage of party organizations' efforts at data campaigning, I used Google News to search for party-focused news coverage across presidential elections from 2012–2020. I took the less traditional approach of using Google News to locate articles that were more likely to have gained popular attention and additional secondary media attention (Ørmen, 2016). While Google News results algorithmically curated nature may mean that search results change, research has shown that to be true of traditional news databases like Factiva and LexisUni as well (Blatchford, 2020). All searches were made manually while logged out of my own Google account to avoid personalized returns.

To create the sample, I engaged both date-constrained and unconstrained searches using the terms "[DNC/RNC]" and data and campaign "from:date [mm/dd/yyyy to:date [mm/dd/yyyy]" in the 6 months leading up to and 3 months following presidential elections (May 1st prior to the election until January 31st following Inauguration Day), and under the "sorted by relevance" setting (Google News' default), for each of the 2012, 2016, and 2020 cycles. Currently, Google sets its archived News results to display a general estimation of total returns to that search (e.g., "about 1,990 results (0.28 seconds)") but only allows users to look at (and click through) the top 50 results. I supplemented this by searching the terms "[DNC/RNC] and data and campaign" without date constraints, but using Google News' "advanced search" button, which is an algorithmically curated search function that purports to return "top news for a given country and language." Collectively, this resulted in 347 articles. After removing results that did not focus on data campaigning or were duplicated across both party searches, a sample of 130 articles remained, and all relevant URLs were entered into a spreadsheet to maintain a stable dataset.

To locate themes and trends within the coverage, I engaged in qualitative thematic analysis (Braun & Clarke, 2006) with attention to how particular narratives and stories about who holds what type of power and skills play out in news stories. Thus, I ground my thematic analysis in questions of "what depictions of the issues around data campaigning, specifically practices of maintenance, modeling, targeting, and testing are circulated to the public?" and "what depictions are missing?" My coding started with deductive attention to categories' presence or absence of campaigns and parties and the types of data practices covered (maintaining, modeling, targeting, testing), and moved into developing inductive codes around which practices were presented in detail, and the details of their description.

4. Findings

4.1. Party and Campaign Data-Work According to Practitioners

Through interviews with political professionals on both the campaign and party sides (including many who have worked in both types of organization), I find that staffers more rooted in party organizations overwhelmingly feel that their work is fundamentally about maintaining data infrastructure, and to a slightly lesser degree, modeling. Campaign staffers and consultants speak much more about targeting and testing, and some also discuss modeling. Importantly, party staffers also discuss how overlooked their core work of maintenance is.

The work discussed by staffers of party organizations, and in fact often ignored by campaign staffers or consultants, is overwhelmingly that of the work of data maintenance. Parties maintain voter files that include but also go beyond contact information such as email addresses and mobile phone numbers, and this work can involve vetting and acquiring new data points, updating existing data, cleaning data, creating original or synthetic data points from other existing data, creating predictive models from any combinations of data, and so on. Moreover, both parties contractually obligate candidates using their files to add any data they procure on their own to their files, obviating that data collection is their purview.

As a leader of a party data team describes it:

> Probably the most expensive and important thing in political data is actually building voter files and aggregating all this data….The voter file system in this country is really, really fragmented…some files haven't been upgraded in years, a lot of it's still there on paper, and there's this third party data that requires sterilization and cleansing process. (Personal communication, 5/27/2019)

Others emphasized the grinding work that this entails, saying "it's not magic, it's just a lot of hard work. It's a lot of cleaning and maintaining, and kind of boring hard work" (personal communication 8/4/2020), and "it's just incredibly time consuming and messy, but one of the most important things is to increase the frequency with which we update data and add more data" (personal communication, 6/24/2020).

Those who did describe maintenance within the campaign explained it as an outlier. One described how most campaigns at all levels below presidential races as being much less concerned with data work of maintenance or modeling, saying:

> When you're building a [campaign] team, the expectation that you're going to *have* data. And like you'll probably hire one person who interfaces with the [outside] data team, and then you're gonna hire most likely firms outside [to handle the data work]. (Personal communication, 8/4/2020)

Another self-proclaimed outlier pointed to their efforts in a presidential primary campaign to obtain data directly from voters and forego outside, non-party data, saying:

> It was really different and innovative. We really could only do it because we had such strong organizing behind us—it was just so many hours from so many people [meaning supporters, not only staffers], I don't know other campaigns that could do it like this. (Personal communication, 11/8/2019)

A staffer with experience in both party and campaign environments described how they saw acquiring data as different from other maintenance practices that many campaigns did not engage in, noting: "We [a presidential campaign] have to clean it up. Most campaigns don't have the manpower for that, and don't do it at all, they just get more emails and use them" (personal communication, 8/4/2020).

The data that parties take so much effort to maintain through new acquisitions, data cleaning, and updating can then be used to create predictive scores that model voting behavior, donor behavior, or persuasive potential. Importantly, staffers from parties and party organizations spoke to both modeling turnout and creating predictive behavioral or persuasion scores. In a recent Medium post, former DNC CTO Nellwyn Thomas touted the party's continued development of additional modeled data points, saying "we provided our users with 75+ unique predictive scores" and a "choice support model, predicting the likelihood a voter supports legal abortions, within 72 hours of the official Dobbs decision" that was passed along to candidates" (Thomas, 2022). Another operative who's worked for the RNC explained to me:

> We also make sure [campaigns] have access to modeling for their races….And not just national, you know, [for] every target Senate race, we do modeling at least once a month right now. In targeted congressional races we also do some. (personal communication, 5/27/2019).

These models tend to be especially useful for down-ballot races, as presidential-level candidates often have their own teams to do this work, though general party work supports all levels of campaigns.

People who worked predominantly in staffer and consultant positions often discussed data as it related to practices of targeting and testing, with much less attention to either modeling or maintaining. Of the few who spoke of maintenance in a campaign environment, their focus was on acquiring data, with one noting: "We bring in contacts–it's our ads and messages that get new phone numbers and emails…that is gold to them [the party organization]" (personal communication, 8/8/2020). Some also discussed modeling, but most of those references were to models they could make use of, not ones they continued to develop (personal communication, 8/14/2019). Another staffer who did describe how they'd done robust, ongoing modeling within a campaign noted how they thought that was an outlier, saying "that's probably only the case within a presidential. And even then, probably not all of them" (personal communication, 8/4/2020).

Overwhelmingly, the campaign and consulting professionals talked about ways they use data to target and test, and how they see those practices as key to successful campaigning. One such professional discussed a variety of data practices they thought were important, from "I think that you can learn a lot about life in general from testing and optimization, from landing pages to emails to Facebook (fundraising) ads…take that attitude that you just don't know what will work best to all things" (personal communication, 8/14/2019). Another with more of a staunch consultant background echoed that, "the less that we have to rely on our gut instinct to as to what's going to work and what's not, it makes everyone's life better" (personal communication, 2/19/2019). This consultant went on to articulate how targeting was most useful in how it could be used to advise campaigns broadly on their overarching messaging strategy, rather than get into highly specific and differentiated targets:

> We can find the people who might be receptive to our messages, which means we can, and we know a little bit more about how better to appeal to them, and how to deliver those messages to them more effectively and efficiently and cost effectively. (Personal communication, 2/19/2019)

One consultant who worked on major Republican campaigns emphasized that some of their best experiences with other data firms dealt with much of the maintaining and modeling, and the campaign "took their memos and their briefings and turn them into ads….I mean a *lot* of ads, so we could test them in Facebook." Another focused on their central focus of using data to target audiences for ads, be they on social media, mobile web ads, or over-the-top TV ads (personal communication, 6/10/2019).

In their considerations of what types of data campaigning they felt received attention from the press, staffers from all types of organizations were clear about what they saw as a dearth of coverage devoted to data maintenance, and about the primacy of targeting and testing. Data staffers on both sides of the aisle chalk this up to the actual work of data maintenance being hard to turn into a story due to its complexity: "[Maintenance] doesn't get a lot of coverage. But you know, I wouldn't necessarily say that's unfair, I don't talk about it a ton either. It's just, it's really, it's complicated. It's convoluted to explain" (personal communication, 7/21/2020). This idea that news norms of novelty and exciting practices made it difficult to write about important, but mundane data work was common. One RNC staffer said that "it's also not sexy, like, how do you sell like, hey, [reporters], we're gonna hire more people to cleanse data!" (personal communication, 5/27/2019).

Some political professionals explained how reporters' expertise and focus on the political side of things, rather than the technology side led to difficulties in gaining coverage: "For the most part, we're talking to political reporters. So we don't even have necessarily the benefit of like, data reporters or more technically savvy reporters who like are going to be breaking down the technology perspective" (personal communication, 6/24/2020). Another explained that "I think that most people [reading the news] don't have a great grasp of data and tech and that makes it hard for journalists to write about" (personal communication, 8/14/2019).

Other data professionals argued that the focus on novelty led to particular data practices getting outsized media attention. One campaign staffer and consultant noted how journalists were ready to cover stories about microtargeting much more often than any other data practices, saying "reporters all the time, they're always looking for me to talk about microtargeting, and I'm just like, unwilling to engage" (personal communication, 8/14/2019). A staffer from the GOP ecosystem described how hard it was to earn coverage making realistic claims: "What I'd like to see, frankly, is less concentration on 'we've found the next big thing' than just the basic blocking and tackling of how data and analytics has been used" (personal communication, 6/10/2019). One democratic data professional described this by saying that "there's definitely not a correlation between the things that are most impactful, and the things that reporters think are really sexy" (personal communication, 7/2/2020).

4.2. Party Data-Work According to News Coverage

If political professionals themselves describe the data work done by national party organizations as substantively different from that which campaigns undertake, a related question remains: How does the press cover the data work done by campaigns and parties? Particularly, if we know that political professionals dedicated to party work see their work of maintenance and modeling as unique, how does that work get covered? Three trends of coverage emerge: (a) an emphasis on campaign work, which largely centers the work of targeting and testing; (b) attention to party and campaign work, where party practices are simplified into "support" of providing data; (c) attention to the details of what goes into maintaining and modeling that has focused on both party work and campaign work. All of these types of coverage are

present, though the former types of coverage are more prevalent than the latter, with overall very little attention to the specificities of party work of maintaining and modeling, especially maintaining.

Overall, even when searching for party-focused stories, for every year's search and on both sides of the aisle, many articles were actually focused on the candidates and their campaign apparatus. Eighteen percent of articles did not mention the RNC or DNC in any way related to data campaigning (often a reference to party fundraising or polling, or a source speaking positively about the candidate was the only mention). Thirty percent of articles center on the campaign, giving only very cursory mention of the party organization's data work. Twenty-eight percent center on the party organization's work. The remaining amount—roughly 40% of the articles—discuss the work of both. The campaign focus was especially true for major news coverage and in-depth reporting: *Bloomberg*'s "Inside the Trump Bunker With 12 Days to Go" (Green & Issenberg, 2016), *MIT Technology Review*'s multi-volume "How Obama's Team Used Big Data to Rally Voters" (Issenberg, 2012), and *Atlantic*'s "Obama's Edge, the Ground Game That Put Him Over the Top" (Ball, 2012) all contain brief mentions of the party organizations, but the main focus is on the campaign. Moreover, in each of those articles, the focus is on practices of targeting and testing. As a result of that slippage, and the overall campaign-focused emphasis on targeting and testing, parties are also portrayed as deeply connected to the work of targeting and message testing.

A second trend of coverage does actually attend to party activity, but flattens and simplifies the work they do—casting the work of maintenance and modeling as something akin to "providing access to data" without describing the degree of work, ongoing attention, and labor that goes into maintaining that data that can be provided. One such article downplays the role of party data, noting that the Trump campaign's so-called "data-push" "also includes commercial data obtained from the RNC and other sources" (Vogel & Samuelsohn, 2016), when the party was the center of the work. Often coverage like this does emphasize organizations' importance in data work, noting that it is central to campaigns' success, but positions the parties as holders, gatekeepers, and suppliers of data, rather than the people who actually procure, access, aggregate, clean, manipulate, and create data. As NBC News wrote, "The presumptive Republican nominee has done little to build its data force and is relying on the Republican National Committee to pick up the bulk of the responsibility for the critical component of its campaign" (Caldwell, 2016).

One limit of this type of coverage is that it often positions data sharing as a single moment of exchanging an asset, rather than a constant process that requires continued attention both inside and outside of election seasons. Examples like "The RNC would offer Trump its full backing, with a promise to share data" (Murray et al., 2018), or "The RNC agreed to allow Trump to use the party's voter file…containing information on more than 200 million Americans" (Vogel & Samuelsohn, 2016) engage in this style of coverage. Such reportage can even describe data in detail, as this example from the *LA Times* does: "the Democratic National Committee's database, containing voting history and demographic information, as well as feedback from contacts with individual voters going back to 1992" (Parsons & Hennessey, 2012), but does not talk about the work processing, wrangling, or otherwise making use of data. Overall, this lack of attention to the work that goes into making data shareable by campaigns across the country, positions data as a stable asset, rather than one that takes labor to constantly construct and maintain.

This coverage also often references the importance of practices like modeling, instead of maintenance work, covering "advances in voter modeling and its heavy investment in the party's ground game" (Vogel, 2016),

and often focusing on simply how much money parties are spending on these tools and assertions of their importance: "The party spent more than $175 million over the past four years to improve predictive modeling to prevent the kind of defeat Mitt Romney faced" (Rosche, 2016). As an outlier that proves a broader norm, one article in *Ad Age* did cover party work related to targeting, but specifically highlighted that it was an anomaly that a party would engage in such activity at a large scale (Kaye, 2016). In this type of coverage, dedication and investment in modelling is referenced more than maintaining.

Finally, articles do occasionally get into the details of what goes into both maintaining and modelling. One 2012 article in Salon, explaining the important development in DNC infrastructure said that the innovation was that "DNC would effectively borrow their files, help clean them up, add new data like donor information and commercially available phone numbers, and then return them for the state party's use." *RollCall* covered of equivalent work by the RNC has noted "The Republican National Committee says its database and models that assign voter scores as well as track how voters react to political advertising and other messages are superior" (Ratnam, 2020). *Forbes* has also covered developments in data work on the RNC side, detailing how the party "developed a data platform that allows candidates to *read and write* data to and from the platform. That means, because they allow access through APIs (routines, protocols and tools to access the data) the Data Center will continuously serve up the latest information for GOP candidates" (Fidelman, 2015).

Following the 2016 election in particular, a sizeable amount of the coverage collected did give detailed insight into the RNC's data campaigning operations—which simply was not the case in other years, or for the DNC in any years studied. Numerous articles in the Fall leading up to the Presidential election centered on how the RNC was supplementing the Trump campaign (Booker, 2016), even specifying that this was "unprecedented" (Kaye, 2016). The RNC was also the subject of so-called "victory lap" articles. One such article, headlined "Republican Party Leaders Take Victory Lap" noted the work of the campaign, consultants, and the party, while quoting RNC Chief of Staff Katie Walsh repeatedly and centering party work (Kamisar, 2016). One Politico article covering an RNC-led briefing the Friday before election day engaged in a meta-commentary that highlighted the fact that the party taking center stage was somewhat abnormal. It specifically noted that:

> The briefing was called ostensibly to highlight the RNC's advances in voter modeling and its heavy investment in the party's ground game. But it also seemed at least partly intended to prove that the Republican Party gave Trump—and all of its 2016 candidates—the tools to succeed in 2016. (Vogel, 2016)

One other trend within detailed coverage is that it emerges when infrastructure breaks. WIRED published an extremely detailed accounting of the DNC's "crumbling" Vertica platform in detail, explaining the maintenance work needed to keep it working, its limitations, and how party organization data staffers were deployed to build a new data infrastructure in its place (Lapowsky, 2019). Similarly, following 2012, numerous news outlets covered the RNC's "autopsy" report, in which the party detailed how it was behind in data infrastructure (Schaeffer, 2013; Wheaton & Shear, 2013). Campaigns' roles in data and infrastructure failures were a key element of post-election accounts of why Clinton did not win (Goldmacher, 2016; Wagner, 2016).

Importantly, the latter two categories of coverage—articles that acknowledge maintaining and modeling and detail its practices—are often heavily rooted in game-frame (Aalberg et al., 2012; Jamieson, 1993) or contest-style coverage. Just as contest-frame coverage of campaigns' is a central part of campaign coverage,

press on parties' data operations often emphasize which party is perceived as "behind" or "ahead" in its data infrastructure and maintenance, as seen in the headline "Democrats Belatedly Launch Operation to Share Information on Voters" (Epstein, 2020). Data campaigning is important enough to garner game frame style coverage of staff changes—firings or additions—was often the hook for press coverage that mostly did not otherwise engage in specifics about data campaigning's specifics. There were stories about RNC firings between cycles (Isenstadt, 2024), as well as coverage of notable departures from top party data roles on both the right and the left (Isenstadt, 2017; Lapowsky, 2019; Nickelsburg, 2019). One staffer on the left pointed to the way that campaigns have had better luck getting the intricacies of data work covered specifically because it can tie into a game frame more directly than party work:

> Early vote data is really good for [getting press coverage]. Like, it's available and actionable, and concrete. You can get journalists to take the bait. You can slice it a bunch of different ways and say look, we're doing well, and they'll cover that. (Personal communication, 8/4/2020)

5. Discussion and Conclusion

Together, these findings paint a picture of party organizations' work of data maintenance–indeed, their primary role, and a focus that is relatively unique to the party organization—as under-recognized, often due to a focus on other DDC practices of modeling, targeting, and testing. While campaigns are likely to engage in particular practices of data acquisition, additional "maintenance work" of cleaning, standardizing data, measuring its importance to campaigns, and so on, are left to party organizations. Data staffers from both party organizations and the campaign/consulting side recognize the difficulty of getting these stories covered, and are very aware of how news norms of covering novel practices and an aversion to explaining complex technological processes hinder coverage. While analysis of news coverage shows that data infrastructure does earn coverage despite data staffers' perceptions that it is ignored completely, there are limitations to that coverage as well. Such coverage often presents party work as the simple task of "providing access" or "handing over" data to campaigns, rather than emphasizing the ongoing work needed to maintain data. News coverage focused on party data efforts does highlight modeling, especially framing predictive work as the key to winning elections, and as a more innovative practice than data maintenance. Moreso, however, news coverage focused on campaign efforts and emphasized targeting and testing practices. Overall, this paints a picture of US DDC scholarship as underplaying the importance of parties.

In many ways, inattention to maintenance is not new, and this work speaks not only to studies of digital political campaigning, but broader work in the "infrastructural turn" in digital media and internet studies (Hesmondhalgh, 2021). As Vinsel and Russell (2020) have argued, maintenance, is key to success, as "maintenance consists of activities that, when done correctly, ensure longevity and sustainability…no innovation can persist without maintenance" (p. 158). Attention to practices of maintenance, rather than just innovation, is important if we want to give a realistic portrayal of emerging technological practice. Revealing these practices also holds instrumental stakes for less professionalized political actors trying to engage in what they assume are key "data-driven practices," but will be superficial and less effective if they overlook the work of maintenance.

This instrumental concern is amplified when we also consider that while maintaining data infrastructure ought to be a fundamental concern of party organizations, these are cash-strapped organizations, and that

maintenance work requires investment in both capital and labor. Party organizations themselves are led by political professionals who generally lack a data background and likely carry a range of assumptions and opinions about data work that may rely on perceptions of what type of data work matters. Recent years have seen the broader campaign technology ecosystem inundated with non-expert elites whose pet projects have failed, like Reid Hoffman's $35 million investment in Alloy, which hoped to bolster data efforts on the left, but ran into political and technical problems alike (Schleifer, 2020). Moreover, maintenance studies have argued that because maintenance work often has lower "occupational prestige" (Vinsel & Russell, 2020)—garnering attention and accolades for certain types of work, while others go unnoticed—it can be harder to draw people to these important jobs. Relatedly, as Kreiss et al. (2020) have shown how women working in data and digital campaigning struggle in a field dominated by a "boys club" culture, does a failure to focus on the maintenance work of party organizations fail to give occupational prestige to particular types of voices?

Although this work focuses on the US, by bringing a focus on party organizations, I hope to enable more robust comparative efforts. As party-centered database developments in the non-US context have emerged as central to DDC work (Anstead, 2017; Dommett, Kefford, & Kruschinski, 2024; Kefford et al., 2023; Munroe & Munroe, 2018) but remain incidental at best to US cases, this article aims to enable more directly comparative questions like do parties engage in practices of data maintenance differently across national contexts, and what are barriers to maintenance across international contexts? This seems especially important as non-US cases are often portrayed as lagging behind the US, but this article's focus on the more mundane maintenance work in the US brings to light some places where similar, not diminished, activity is actually taking place. Relatedly, this contributes to demystifying practices of DDC across international contexts, pointing to the mundane practices of data cleaning and systematization that are central to the work, rather than positioning it as a radical new invention.

By bringing together the direct accounts of political professionals with news coverage, this study aims to investigate both what party organization practices look like and how they are covered. Still, both methods have limitations. The political professionals I spoke with are all at the highest level of the field, often running teams in presidential races and leading data teams at party organizations or working at the top data consultancies. While this certainly makes them experts, it also means their experience is quite different than those who do data work at down-ballot races. Even though such races may be more reliant on party data work (at the national and state level), that work is likely to be even more invisible, as it is all done off-site, and local staffers are unlikely to have close ties with or dedicated staff who can speak to data work. In assessing news coverage, while the number of news stories was relatively small (n = 130), the goals here were not to give insight into things like frequency of coverage. Instead, this work unearths broad themes that are present within the coverage that does exist. Future inquiry into more deductive assessments of how often and under what conditions various types of data campaigning are covered is warranted. Whether maintenance work continues to be ignored in contexts where more attention is paid to party work is an interesting future avenue for research.

Despite these limits, this article offers insight into the role of party organizations' particular types of data work in the US and is a first step toward clearer comparative analysis with party-centering systems. In doing so, I find that parties in the US engage in robust, and indeed deeply important, practices of DDC that are fundamentally different from those of candidate-centered campaigns. Moreover, these party activities are

widely overlooked and oversimplified in the news, due to the unique favoring of campaigns and the difficulty of covering mundane maintenance work.

Conflict of Interests

The author has no conflicts of interest.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Aalberg, T., Strömbäck, J., & de Vreese, C. H. (2012). The framing of politics as strategy and game: A review of concepts, operationalizations and key findings. *Journalism*, *13*(2), 162–178.

Anstead, N. (2017). Data-driven campaigning in the 2015 United Kingdom general election. *The International Journal of Press/Politics*, *22*(3), 294–313.

Arbour, B. (2014). *Candidate-centered campaigns: Political messages, winning personalities, and personal appeals*. Palgrave Macmillan.

Baldwin-Philippi, J. (2017). The myths of data-driven campaigning. *Political Communication*, *34*(4), 627–633.

Baldwin-Philippi, J. (2019). Data campaigning: Between empirics and assumptions. *Internet Policy Review*, *8*(4). https://doi.org/10.14763/2019.4.1437

Ball, M. (2012, October 24). Obama's edge: The ground game that could put him over the top. *The Atlantic*. https://www.theatlantic.com/politics/archive/2012/10/obamas-edge-the-ground-game-that-could-put-him-over-the-top/264031

Blatchford, A. (2020). Searching for online news content: The challenges and decisions. *Communication Research and Practice*, *6*(2), 143–156.

Booker, B. (2016, September 2). RNC hires staff to supplement Trump's lacking ground game in critical states. *NPR*. https://www.npr.org/2016/09/02/492434914/rnc-hires-staff-to-supplement-trumps-lacking-ground-game-in-critical-states

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Caldwell, L. A. (2016, May 31). Donald Trump's campaign behind on critical data operation. *NBC News*. https://www.nbcnews.com/politics/2016-election/donald-trump-s-campaign-far-behind-critical-data-operation-n583156

Cohen, M., Karol, D., Noel, H., & Zaller, J. (2008). *The party decides: Presidential nominations before and after reform*. University of Chicago Press.

Dommett, K., Barclay, A., & Gibson, R. (2024). Just what is data-driven campaigning? A systematic review. *Information, Communication & Society*, *27*(1), 1–22.

Dommett, K., Kefford, G., & Kruschinski, S. (2024). *Data-driven campaigning and political parties: Five advanced democracies compared*. Oxford University Press.

Dommett, K., Kefford, G., & Power, S. (2021). The digital ecosystem: The new politics of party organization in parliamentary democracies. *Party Politics*, *27*(5), 847–857.

Dwyre, D. (2010). A work in progress: Parties and democracy in the United States. In K. Lawson (Ed.), *Political parties and democracy: The Americas—Volume 1* (pp. 27–47). Praeger.

Epstein, R. (2020, September 6). Democrats belatedly launch operation to share information on voters. *The New York Times*. https://www.nytimes.com/2020/09/06/us/politics/Presidential-election-voting-Democrats.html

Fidelman, M. (2015, August 3). Exclusive: Republicans launch game-changing data center that will forever change politics. *Forbes*. https://www.forbes.com/sites/markfidelman/2015/08/03/exclusive-republicans-launch-game-changing-data-center-that-will-forever-change-politics

Galvin, D. J. (2012). The transformation of political institutions: Investments in institutional resources and gradual change in the national party committees. *Studies in American Political Development*, *26*(1), 50–70.

Gerken, H., & Fishkin, J. (2014). The two trends that matter for party politics. *NYU Law Review Online*, *89*, 32–47.

Gibson, R. (2020). *When the nerds go marching in: How digital technology moved from the margins to the mainstream of political campaigns*. Oxford University Press.

Goldmacher, S. (2016, September 7). *Hillary Clinton's 'invisible guiding hand.'* POLITICO Magazine. https://www.politico.com/magazine/story/2016/09/hillary-clinton-data-campaign-elan-kriegel-214215

Green, J., & Issenberg, S. (2016, October 27). Inside the Trump bunker, with days to go. *Bloomberg*. https://www.bloomberg.com/news/articles/2016-10-27/inside-the-trump-bunker-with-12-days-to-go

Hersh, E. (2015). *Hacking the electorate: How campaigns perceive voters*. Cambridge University Press.

Hesmondhalgh, D. (2021). The infrastructural turn in media and internet research. In P. McDonald (Ed.), *The Routledge companion to media industries* (pp. 132–142). Routledge.

Isenstadt, A. (2017, August 28). *Another RNC data guru heads for the exits*. POLITICO. https://www.politico.com/story/2017/08/28/conor-mcguire-leaves-republican-national-committee-242099

Isenstadt, A. (2024, March 11). *Bloodbath at RNC: Trump team slashes staff at committee*. POLITICO. https://www.politico.com/news/2024/03/11/bloodbath-at-rnc-trump-team-slashes-staff-at-committee-00146368

Issenberg, S. (2012, December 19). *How Obama's team used big data to rally voters*. MIT Technology Review. https://www.technologyreview.com/s/509026/how-obamas-team-used-big-data-to-rally-voters

Jamieson, K. H. (1993). *Dirty politics: Deception, distraction, and democracy*. Oxford University Press.

Kamisar, B. (2016, November 14). Republican Party leaders take victory lap. *The Hill*. https://thehill.com/blogs/ballot-box/presidential-races/305922-republican-party-leaders-take-victory-lap

Karpf, D. (2016). *Analytic activism: Digital listening and the new political strategy*. Oxford University Press.

Kaye, K. (2016, May 20). *RNC reserves an unprecedented $150 million in digital video ads*. Ad Age. https://adage.com/article/campaign-trail/rnc-makes-unprecedented-150-million-digital-ad-buy/304107

Kefford, G. (2021). *Political parties and campaigning in Australia: Data, digital and field*. Springer.

Kefford, G., Dommett, K., Baldwin-Philippi, J., Bannerman, S., Dobber, T., Kruschinski, S., Kruikemeier, S., & Rzepecki, E. (2023). Data-driven campaigning and democratic disruption: Evidence from six advanced democracies. *Party Politics*, *29*(3), 448–462.

Kreiss, D. (2016). *Prototype politics: Technology-intensive campaigning and the data of democracy*. Oxford University Press.

Kreiss, D., Adams, K., Ciesielski, J., Fernandez, H., Frauenfelder, K., Lowe, B., & Micchia, G. (2020). *Recoding the boys' club: The experiences and future of women in political technology*. Oxford University Press.

Kruikemeier, S., Vermeer, S., Metoui, N., Dobber, T., & Zarouali, B. (2022). (Tar)getting you: The use of online political targeted messages on Facebook. *Big Data & Society*, *9*(2). https://doi.org/10.1177/20539517221089626

Lapowsky, I. (2019, April 2). *Inside the Democrats' plan to fix their crumbling data operation*. WIRED. https://www.wired.com/story/democrats-fix-crumbling-data-operation

Masket, S., & Noel, H. (2021). *Political parties*. W. W. Norton.

Munroe, K. B., & Munroe, H. D. (2018). Constituency campaigning in the age of data. *Canadian Journal of Political Science/Revue Canadienne de Science Politique*, *51*(1), 135–154.

Murray, S., Reston, M., Bash, D., & Perez, E. (2018, March 21). *Inside the Trump campaign's ties with Cambridge Analytica*. CNN Politics. https://www.cnn.com/2018/03/21/politics/trump-campaign-cambridge-analytica/index.html

Nickelsburg, M. (2019, July 29). *DNC's new tech leader talks about what went wrong in 2016 and how Dems are preparing for 2020*. GeekWire. https://www.geekwire.com/2019/dncs-new-tech-leader-talks-went-wrong-2016-dems-preparing-2020

Nickerson, D. W., & Rogers, T. (2010). Do you have a voting plan? Implementation intentions, voter turnout, and organic plan making. *Psychological Science*, *21*(2), 194–199.

Nickerson, D. W., & Rogers, T. (2014). Political campaigns and big data. *The Journal of Economic Perspectives*, *28*(2), 51–73.

Nielsen, R. K. (2012). *Ground wars: Personalized communication in political campaigns*. Princeton University Press.

Ørmen, J. (2016). Googling the news: Opportunities and challenges in studying news events through Google Search. *Digital Journalism*, *4*(1), 107–124. https://doi.org/10.1080/21670811.2015.1093272

Owen, D. (2013). Political Parties and the Media: The Parties Respond to Technological Innovation. In M. D. Brewer (Ed.), *The parties respond* (5th ed., pp. 237-264). Routledge.

Parsons, C., & Hennessey, K. (2012, November 13). Obama campaign's investment in data crunching paid off. *Los Angeles Times*. https://www.latimes.com/archives/la-xpm-2012-nov-13-la-na-obama-analytics-20121113-story.html

Ratnam, G. (2020, July 29). Republicans and Democrats battle over who has better voter data effort. *Roll Call*. https://rollcall.com/2020/07/29/republicans-and-democrats-battle-over-who-has-better-voter-data-effort

Roemmele, A., & Gibson, R. (2020). Scientific and subversive: The two faces of the fourth era of political campaigning. *New Media & Society*, *22*(4), 595–610.

Rosche, J. (2016, November 2). *RNC chief of staff Katie Walsh describes the $175 million tool to win elections*. CNN Politics. https://www.cnn.com/2016/11/02/politics/katie-walsh-party-people-podcast/index.html

Russell, A., & Vinsel. L. (2016, April 7). Hail the maintainers. *Aeon*. https://aeon.co/essays/innovation-is-overvalued-maintenance-often-matters-more

Schaeffer, A. (2013, March 22). The devil is in the data. *US News & World Report*. https://www.usnews.com/opinion/articles/2013/03/22/whats-missing-from-the-rncs-2012-autopsy

Schleifer, T. (2020, December 15). *Inside the implosion of the $35 million startup meant to fix the Democratic Party*. Vox. https://www.vox.com/recode/22175186/alloy-drama-reid-hoffman-democratic-data-startup

Stromer-Galley, J. (2014). *Presidential campaigning in the internet age*. Oxford University Press.

Thomas, N. (2022, December 1). *Reflecting on DNC Tech in 2022*. DNC Tech Team. https://medium.com/democratictech/reflecting-on-dnc-tech-in-2022-597287f10d67

Vinsel, L., & Russell, A. (2020). *The innovation delusion*. Penguin Random House.

Vogel, K. P. (2016, November 7). *RNC model showed Trump losing*. POLITICO. https://www.politico.com/story/2016/11/rnc-model-showed-trump-losing-231074

Vogel, K. P., & Samuelsohn, D. (2016, January 5). Trump quietly builds a data juggernaut. *POLITICO*. https://www.politico.com/story/2016/01/trump-builds-data-juggernaut-217391

Wagner, J. (2016, November 9). Clinton's data-driven campaign relied heavily on an algorithm named Ada. What didn't she see? *The Washington Post*. https://www.washingtonpost.com/news/post-politics/wp/2016/11/09/clintons-data-driven-campaign-relied-heavily-on-an-algorithm-named-ada-what-didnt-she-see

Wheaton, S., & Shear, M. D. (2013, March 18). Blunt report says G.O.P. needs to regroup for '16. *The New York Times*. https://www.nytimes.com/2013/03/19/us/politics/republicans-plan-overhaul-for-2016-primary-season.html

About the Author

Jessica Baldwin-Philippi is an associate professor in Fordham University's Communication and Media Studies Department, and an expert in the study of digital campaigning. She has been published in the journals *Political Communication*, *New Media & Society*, *International Journal of Communication*, among others.