ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Darius, Philipp et al.

Article

Election research in the age of regulated data access under the EU Digital Services Act

Internet Policy Review

Provided in Cooperation with:
Alexander von Humboldt Institute for Internet and Society (HIIG), Berlin

Suggested Citation: Darius, Philipp et al. (2026) : Election research in the age of regulated data access under the EU Digital Services Act, Internet Policy Review, ISSN 2197-6775, Alexander von Humboldt Institute for Internet and Society, Berlin, Vol. 15, Iss. 1, pp. 1-25, https://doi.org/10.14763/2026.1.2080

This Version is available at:
https://hdl.handle.net/10419/339537

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

https://creativecommons.org/licenses/by/3.0/de/deed.en

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

INTERNET POLICY REVIEW

Volume 15 **Issue** 1

RESEARCH ARTICLE

OPEN ACCESS

PEER REVIEWED

Election research in the age of regulated data access under the EU Digital Services Act

Philipp Darius *Hertie School*
Johannes Breuer *Center for Advanced Internet Studies (CAIS)*
Simon Kruschinski *GESIS - Leibniz Institute for the Social Sciences*
simon.kruschinski@uni-mainz.de
Felicia Loecherbach *University of Amsterdam* f.loecherbach@uva.nl
Jasmin Riedl *University of the Bundeswehr Munich*
Sebastian Stier *GESIS - Leibniz Institute for the Social Sciences*

DOI: https://doi.org/10.14763/2026.1.2080

Published: 16 February 2026
Received: 18 February 2025 **Accepted:** 27 June 2025

Funding: The authors have received no additional funding for this research.
Competing Interests: The author has declared that no competing interests exist that have influenced the text.
Licence: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License (Germany) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. https://creativecommons.org/licenses/by/3.0/de/deed.en
Copyright remains with the author(s).

Citation: Darius, P., Breuer, J., Kruschinski, S., Loecherbach, F., Riedl, J., & Stier, S. (2026). Election research in the age of regulated data access under the EU Digital Services Act. *Internet Policy Review*, *15*(1). https://doi.org/10.14763/2026.1.2080

Keywords: Digital Services Act (DSA), Regulatory governance, Digital platforms, Election research, Data access

Abstract: Political debates, campaigns, and advertising increasingly take place on online platforms. However, research on election communication has been significantly hampered as formerly accessible data sources have been withdrawn or commercialised. With the implementation of the EU Digital Services Act (DSA) a number of platforms have (re-)established data access modalities for 'public' data via APIs or specific portals. How access to public data is implemented, what it contains, and who gets access all depends on decisions by the platforms. This leads to a series of inconsistencies, challenges, and limitations for election research. In this contribution, we discuss the implications of regulated data access under the DSA for election research. We first review central research questions and relevant data types for election research. Then, we provide a historical overview of how data access modalities have changed over the last two decades. Next, we discuss relevant articles of the DSA that aim to improve data access for academic research as well as different data access paths and modalities, including alternatives to APIs such as web scraping and data donations. Finally we summarise key challenges and formulate requirements for data access to enable robust and reproducible election research.

Introduction

In many countries the production and consumption of political information has significantly shifted to online platforms, typically concentrating on a few major social media platforms. Correspondingly, political parties and candidates use social media as communication channels, especially for election communication and campaigning. Researching electoral communication or, more generally, political behaviour on digital platforms, however, is often hampered by changing data access modalities and a lack of transparency of digital platforms. In recent years, the closure of APIs by Facebook and Twitter/X and failure of the industry-academic collaboration Social Science One led researchers to diagnose an "APIcalypse" (Bruns, 2019), a "data abyss" (De Vreese and Tromble, 2023), and a "post-API age" (Freelon, 2018). However, both the regulatory environment in various countries as well as ownership and policies of major platforms like X have been shifting fast. A major recent legislative change with substantial impact for data access for academic research is the Digital Services Act (DSA) in Europe, which was introduced in February 2024. It mandates access to public data from Very Large Online Platforms (VLOPs) as well as extended data access to non-public data for vetted researchers. We argue that the DSA (and potentially also similar legislation in other jurisdictions) introduces a new phase in data access for election research on digital platforms that we call the "regulated data access age". In this article, we describe the promises of this regulated data access age for election research, identify remaining challenges, and formulate requirements for improving data access conditions for election research[1].

Election research and online platforms

Elections constitute one of the most pivotal moments in democratic systems, marked by heightened levels of political communication as political actors and citizens alike engage intensively in public discourse. In recent years, this discourse has increasingly migrated to social media platforms, fundamentally reshaping the dynamics of electoral communication. We focus on election research because online platforms, and the data they hold, have become indispensable to the study of contemporary electoral processes. In addition, the field of election research addresses key questions related to "systemic risks" within the European Union and beyond, which lie at the core of the Digital Services Act (DSA) data access frame-

1. We use the more general term election research instead of "online election research" or "digital election research" as online platforms have become so ingrained into communication and interactions around and about elections that, nowadays, it is impossible to conduct a comprehensive analysis of elections without considering the online/digital sphere.

work. The following sections outline the conceptual, methodological, and regulatory implications for election research, followed by an overview of existing research on the supply and demand sides of online political behaviour.

The overall goal of analysing data from digital platforms for election research is to understand and explain online political behaviour by citizens and political actors like parties and candidates (also often referred to as political elites). The former is typically referred to as the "demand side" (i.e., citizens looking for information or engaging with political content), while the latter is commonly described as the "supply side" (i.e., political actors disseminating content and communicating with the public or other actors; see van Aelst et al., 2017; Xenos et al., 2017). In practice, phenomena like digital publics, online platforms, and social media as digital campaign tools can collapse the classical supply-demand-side division of political communication as they enable direct interaction between politicians, journalists and politically interested citizens (Gibson, 2015; Jungherr et al., 2020; Stier et al., 2018). From an analytical perspective, however, the distinction is still relevant, especially for design choices in empirical studies on political communication as well as the technical, legal, and ethical questions that are associated with them. For our discussion of data access regimes and their implications, we want to focus particularly on the preconditions for research studying how politicians, parties, and other political actors use social media and other online platforms for communication and campaigning (e.g., Haßler et al., 2021; Kruschinski & Bene, 2022; Votta et al., 2024a).

Another relevant distinction classifies the research designs that are needed to collect digital behavioural data into platform-centred and user-centred data collection approaches (Wagner et al., 2025). In platform-centred designs, researchers sample data from specific platforms based on the research question(s) at hand. This sampling can, e.g., be based on (combinations of) features, such as users, topics, hashtags, search queries or time (see Kruschinski et al., 2024; Müller et al. 2022; Riedl et al. 2024 for some examples). The data is then typically directly collected from the respective platforms. By contrast, in user-centred designs (see, e.g., Breuer et al., 2023a; Halavais, 2019), researchers first recruit participants[2] and those are then actively involved in collecting or providing the data.

Besides these important methodological distinctions, from a more substantive perspective, the use of social media and other online platforms for communication by

2. See Breuer et al. (2023b) for a discussion of whether or in what cases "participants" is the right term for individuals whose data are being used in studies with digital behavioural data collected via platform APIs.

political actors is associated with both opportunities and risks (or has positive and negative aspects; see, e.g., Roemmele et al., 2020). or example, the engagement driven structures of some platforms tend to algorithmically favour negative, radical and extreme political content and are, hence, considered to strengthen radical populist parties (Bouchaud et al., 2023; Huszár, 2022; but see Kruschinski et al., 2024). Recently, with the rise of generative AI and its increasing entanglement with digital platforms and their algorithms[3]researchers have been raising concerns about its harmful potential as an instrument for producing and disseminating disinformation about the electoral process and the involved candidates (Brautović et. al. 2024; Corsi et. al., 2024; Darius & Roemmele, 2023; Kruschinski & Votta, 2025; Lewandowsky et al., 2023). On the other hand, social media have extended the toolbox of campaign tools for candidates, politicians and parties (Dommett et al., 2024; Gibson, 2015; Jungherr 2016; Kruschinski & Bene, 2022; Votta et al., 2024a) and, thus, increased opportunities for communication and interaction with (potential) voters.

Despite the possibilities and opportunities for communication and interaction between political parties and politicians and (potential) voters, much of the research on elections and online platforms has focused on risks and negative aspects. The predominance of research on social media's risks and downsides can be traced to a confluence of factors. Real-world events (like radical or populist parties` electoral success and online misinformation campaigns) have shattered earlier optimism and created a climate of concern that academia has responded to. One pivotal moment was the 2016 election of Donald Trump in the US, alongside the Brexit referendum. These unexpected outcomes prompted observers to seek explanations in digital campaigning. Bots, foreign interference, targeted ads, echo chambers, and disinformation were blamed for influencing voters.

Another important event that also had consequences for data access for research itself is the Facebook/Cambridge Analytica scandal (see Venturini & Rogers, 2019). This scandal involved the unauthorised harvesting of personal data from up to 87 million Facebook users by the British political consulting firm Cambridge Analytica in the mid-2010s with the goal of using it to influence election results. The result of this was that access to user data via the Facebook Graph API was closed.

Regardless of whether studies look at the opportunities or the risks associated with the use of online platforms for political communication, a key necessity for

3. Two prominent recent examples are the integration of Meta AI in WhatsApp or the chatbot Grok on X.

studying such topics is access to relevant data. Without reliable access to high-quality data, it is difficult or even impossible to establish a solid empirical foundation needed for informing policies to counter potentially adverse effects or harness beneficial ones.

In the following, we will discuss research on the supply and demand to showcase what empirical knowledge can be derived from digital behavioural data for election research.

2.1. Researching the supply side: elite communication on online platforms

Election research on the supply side looks at how political elites use online platforms to communicate and interact. Many of the studies in this area concentrate on campaign behaviour as a period in which political actors actively seek to influence citizens' opinions and decisions (e.g., Norris, 2000; Schmitt-Beck, 2000). In general, online platforms have decreased the gatekeeping power of traditional media and allow activists, partisans, special interest organisations, and also foreign actors to participate in the electoral debate, expressing political beliefs, and seeking to influence public opinion, political agendas and, potentially, also citizens' individual voting decisions (Darius, 2022; Gibson, 2015; Haßler et al., 2021).

Research on the communication of political elites on online platforms is as diverse as the platforms and interactions being studied. Depending on the political system, transnational (European), national, state, regional, and local elections are studied. To account for the complexity of contemporary party competition, research is becoming increasingly specialised, e.g., taking into account regional variation across different countries (Darius et al., 2025; Dommett et al., 2025; Votta et al., 2024a) subnational states (Maier et al., 2023) or intraparty heterogeneity in political positions (Sältzer, 2022). Central research areas are, for example, the salience of political issues and intermedia agenda setting (Barberá et al., 2019; Gilardi et al., 2022; Haßler et al., 2024), negative campaigning (Maier et al., 2023; Nai et al., 2020), how politicians circumvent journalists and directly reach supporters (Jungherr et al., 2019), or to what extent direct communication by politicians increases the political knowledge of citizens (Munger et al., 2022; Popa et al., 2020). Other studies show that the sharing of low-quality information by political elites can be strategic, with a higher exposure among conservative users (Hjorth & Adler-Nissen 2019), a higher prevalence of untrustworthy content among radical right populist parties (Törnberg & Chueri, 2025) and politicians from the Republican Party in the US (Eady et al., 2023; Lasser et al., 2023; Greene, 2024).

Social media data as a data source has significantly advanced research on party politics and election campaigning, as online platforms provide much more data with a higher volume and granularity than other widely used data sources in this field, such as party manifestos, party press releases or parliamentary speeches (Lehmann et al., 2024; Rauh & Schwalbach, 2020). Social media data can include the content of political communication for hundreds or thousands of political actors of interest, including central party accounts and individual candidates (e.g., Haßler et al., 2021; Kruschinski & Bene, 2022; Votta et al., 2024a). In addition, the data is produced at a granular temporal resolution, allowing for nowcasting (Kruschinski et al., 2025; Müller et al., 2022), longitudinal analysis (Bruns et al., 2021; Ceron et al., 2016; Votta et al., 2024a), or the analysis of strategic communication as a reaction to external political shocks, such as unfolding campaign events.

A crucial part of election campaigning and strategic political communication online happens through (targeted) advertising (Dommett et al, 2025; Votta et al., 2024a). Research on political advertising on social media has primarily focused on its normative implications, deployment, and effects. Several studies indicate that microtargeting is widely adopted within social media advertising in many countries (Kruschinski & Bene, 2022; Votta et al., 2024a), though its sophistication varies across contexts (Dommett et al., 2024; Votta et al., 2024a). Research on its impact suggests that microtargeting has limited effects on vote choice but may reinforce partisan attachment and influence voter attitudes when ads align with personal traits and issue priorities (Chu et al., 2024; Binder et al., 2022; Zarouali et al., 2022). However, studies indicate that, despite increasing regulation, platforms do not successfully check advertisements for false information. A report by the NGO Global Witness and the NYU engineering department, for instance, indicated that most platforms failed to remove ads containing election fraud narratives before elections in most countries and seemed to focus their efforts on the US and major European countries (Global Witness Briefing, 2022). Something that has been rarely studied so far in this context are programmatic ads that are traded via advertisement exchange platforms. These ads enable targeting methods such as very specific targeting of IP addresses like company IPs or based on browsing behaviour data collected by second and third parties that may infringe with the GDPR (Utz et al., 2022). Important questions that remain address the extent to which individual or repeated ads can have a persuasive effect on voters' electoral behaviour, how parties and politicians use generative AI for content production and dissemination, and whether online platforms start to appropriately moderate targeted ads that often allow for the dissemination of content that contains false information and other sorts of harmful content (Bouchaud & Liénard, 2024). These

examples already indicate that it is also important to study the users that receive or are targeted by the online communication of political elites and what effects this may have on them. This perspective is typically taken in research on the demand side

2.2. Researching the demand side: effects of political communication and online behaviour

Generally speaking, research on the demand side looks at how and where citizens are exposed to/receive political information and communication and how it influences their attitudes or behaviours. A significant body of demand-side research looks at the selection and consumption of political information and use, especially news exposure (Guess, 2021; Stier et al., 2025) or how citizens choose which parties, politicians, or political influencers to follow on online platforms (Wojcieszak et al., 2022; Gibson et al., 2023). Selective exposure studies show that voters tend to prefer content that aligns with their predispositions (e.g., populist-minded individuals favor anti-establishment posts, partisans prefer like-minded sources), and this tendency contributes to a fragmented media consumption across electorates (Mangold et al., 2024). Algorithmic curation on platforms might reinforce these patterns by tailoring feeds to user preferences, resulting in more exposure to like-minded sources and ideological segregation (González-Bailón et al., 2023).

Another area of research on the demand side of political online communication is concerned with effects of disinformation and misinformation during election campaigns. More generally, studies have shown that Facebook or WhatsApp use have positive effects on the acquisition of political knowledge during election campaigns (Allcott & Gentzkow, 2020; Ventura et al., 2025). Still, the quality of information matters, as exposure to disinformation and misinformation can also skew voting decisions or preference for a candidate (Bovet & Makse, 2019; Eady et al., 2023; Zimmermann & Kohring, 2020). However, in line with the minimal effects found for linear mass media, even sophisticated field experiments have found no or only very small effects of various facets of digital media use on political attitudes, affective or issue polarisation (Guess et al., 2023).

Notably, supply side and demand side can also be studied together as, e.g., studies on populist radical right parties by Mols and Jetten (2020) and Stier et al. (2025) have shown. Regardless of the perspective or focus, an important source of challenges for election research is that digital platforms underlie constant changes that can relate to ownership, design or technical features. This also means that the conditions for access to online digital platform data have drastically changed in re-

cent years. The DSA promises an improvement of this situation by establishing regulated data access.

From wild west, to post-API age, to regulated data access under the EU DSA

What research on the supply side and the demand side has in common is that both streams of research need reliable access to data from online platforms. Data access modalities have been subject to various changes within the past two decades, in particular with respect to Application Programming Interfaces (APIs) offered by platforms. While there are other methods for collecting online platform data (Breuer et al., 2020), APIs have been widely used, in election research as well as other fields, and are, hence, often at the centre of discussion in academic debates and publications. We will discuss some alternative data access options for election research later on, but focus specifically on the historical development of data access via APIs in this section.

APIs are often called the backbone of the internet since they enable communication and data exchanges between clients (computers/devices/servers). For researchers interested in socio-technical and political mechanisms on digital platforms, institutionalised platform APIs provide a scalable and legally clearly defined method of collecting research data (Breuer et al., 2020). This presumes, however, that the data providers are able and willing to build, maintain and document an appropriate API infrastructure. Considering that most platforms have an interest in developers building programs based on their APIs, many platforms have API infrastructures in place and can allow for extended access (with more tokens/limits) for researchers. While there are various types of APIs, research or developer APIs are commonly Representational State Transfer Application Program Interface (REST APIs), typically implemented using the http(s) client-server protocol applying standard methods, such as "get, post, put, delete". An important thing to keep in mind is that most platform APIs were created for developers and their use for academic research has often just been tolerated without being an intended use case originally. Eventually, several platforms have developed specific APIs or API access programs for researchers but this happened later than the first wave of social science research that made extensive use of platform APIs.

The rise of API use in the social and behavioural sciences started in the early 2010s with projects like MyPersonality that gathered large amounts of user data via the Facebook Graph API to study and predict personality traits and other attributes (Kosinski et al., 2013).[4]Some platforms offered relatively unrestricted access

to user data. Besides Facebook, Twitter, e.g., provided generous and easy to use Streaming and REST APIs. Later on, the platform even offered a dedicated and extended API access track for academic research. These data access opportunities were used in a large number of studies on various topics, including much of the election-related research cited above. To describe these broad data collection opportunities, Tromble (2021) has used the term "Data Golden Age". The first big event that changed this initially rather liberal data access landscape was the restrictions put in place by Facebook in the aftermath of the Cambridge Analytica scandal in 2018. The change of the app review process and partial closure of Facebook's API also resulted in the discontinuation of widely used research tools like Netvizz (Rieder, 2013; 2018). Researchers reacted to these restrictions placed on platform data access by diagnosing an "APIcalypse" (Bruns, 2019) or the dawn of a new "post-API age" for computational research (Freelon, 2018). In essence, the opinions voiced in these articles are that the dependence on data access through APIs that were voluntarily provided by commercial platforms was a risk and that researchers will (increasingly) need to rely on alternative methods for collecting online platform data. Others, however, have also pointed out the positive aspects of such changes, e.g., with regard to data protection and privacy as an end to the 'wild west' of social media research (Puschmann, 2019).

Despite the experiences with closures or severe restrictions of platform APIs and the critical discussions this has sparked, researchers in election research and computational social science have continued to rely on APIs for data access. Recently, however, further closures or severe restrictions of APIs or tools like CrowdTangle[5]have underlined the inherent risks and problems associated with this type of data access, which marks a significant shift in the accessibility of public social media data for researchers (as well as journalists and civil society organisations). A key event in this was the acquisition of Twitter by Elon Musk and its renaming to X that was followed by a closing of the free Academic Research API access.

Figure 1 provides a summary of the key changes within major platforms Twitter/X, Facebook, Instagram, YouTube, and TikTok that inhibited election research as well as relevant regulatory developments (with a focus on the European Union). The figure highlights three phases of data access to platform data and APIs for researchers. The first phase of unregulated and extensive data access is marked as

4. Notably, already in 2009 the sale of Yahoo to Microsoft resulted in a closure of API-based mash-up tools that were used by some researchers for data extraction on search results and other metrics (Rieder, 2009).

5. CrowdTangle, a tool acquired by Facebook in 2016 that allowed users to track the spread and engagement of public content across Facebook, Instagram, and Reddit.

the 'Data Golden Age' by Tromble (2021). Following Freelon (2018), we label the second phase marked by closures or severe restrictions in data access as the 'Post-API Age'. With the introduction of the EU DSA, we see the beginning of a third phase of mandated data access that we would call a 'Regulated API Age'. In a recent preprint, Mimizuki et al. (2025) use the term "Post-Post API Age". However, since we believe that the aspect of regulation is a key feature of this new data access age that also sets it apart from the largely unregulated setting in the "Data Golden Age" (or "Wild West"), we prefer to use the term "Regulated API Age". The key piece of legislation for this regulation is the EU DSA. In the following, we discuss some of the main goals and elements of the EU DSA as well as their implications for election research.

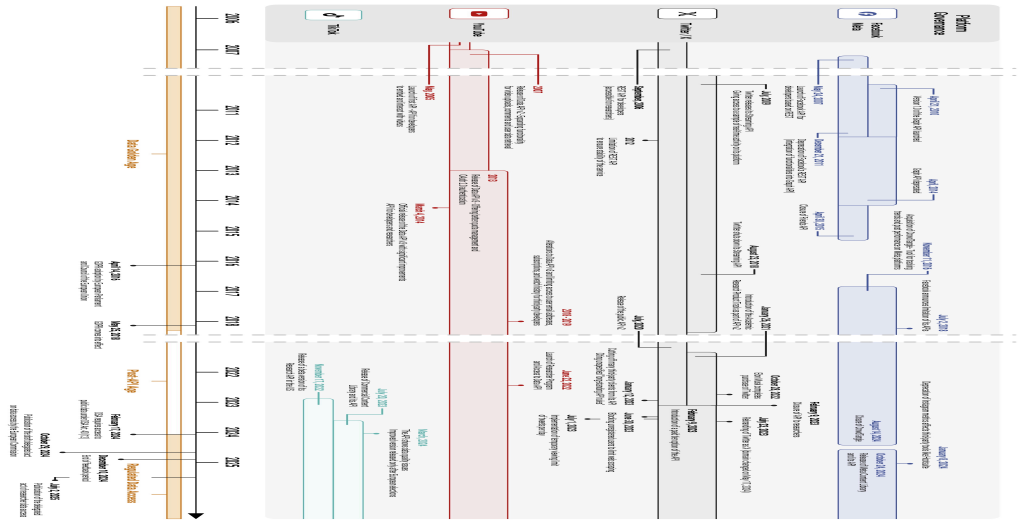

FIGURE 1: Timeline of events affecting platform data access for research.

The EU Digital Services Act and its access modalities for election research

The EU DSA aims to promote safe online environments for European citizens and entails new data access provisions for researchers to investigate and better understand systemic risks that digital platform services may pose to individuals and society as a whole. To enable the investigation and mitigation of systemic risks the DSA includes several data access and transparency components, such as access to public data under DSA 40(12), extended research data access for vetted researchers under 40(4), as well as the mandatory provision of data for the European moderation database under DSA 24(5) and public ad archives under DSA 39. In the following we will discuss these components and their significance for election researchers.

4.1. Access to public data under DSA 40(12)

From February 2024 very large online platforms (VLOPS) and very large search engines (VLOSEs) are mandated under Article 40(12) of the DSA to enable real-time data access for researchers. The text describes data access as follows: "*Providers of very large online platforms or of very large online search engines shall give access without undue delay to data, including, where technically possible, to real-time data, provided that the data is publicly accessible in their online interface by researchers.*"

In their efforts to implement the data access mandates included in the DSA, most social media platforms have introduced or extended data access options for researchers. This has typically been done via dedicated APIs or API access programs. Notably, APIs are only one form of providing such data access and platforms decide themselves which researchers receive access to their API under DSA 40(12). Instead of dedicated research(er) APIs, some major platforms have decided to offer sandbox environments or digital dashboards that avoid an export of user data from the company's server or strictly limit the ability to export data. The Meta Content Library[6]is an example of this that provides both a web interface as well as an API that can only be used in a secure environment that strictly limits data export.

For election research, access to such public data is especially interesting for studies on the supply side investigating how political elites use different platforms to communicate with (potential) voters as well as other actors and organisations.

4.2. Vetted researcher data access under DSA 40(4)

Beyond access to public data provided by the platforms themselves, under 40(4) of the DSA, vetted researchers shall receive access to extended data based on the criteria further specified in the draft delegated act (Digital Services Act, 2022).

"Data related to users, such as profile information, relationship networks, individual-level content exposure and engagement histories; interaction data, such as comments or other engagements; data related to content recommendations, including data used to personalise recommendations; data related to ad targeting and profiling, including cost per click data and other measures of advertising prices; data related to the testing of new features prior to their deployment, including the results of A/B tests; data related to content moderation and governance, such as data on algorithmic or other content moderation systems and processes, archives or repositories documenting moderated

6. Meta Content Library (see https://transparency.meta.com/en-us/researchtools/meta-content-library/)

content, including accounts as well as data related to prices, quantities and characteristics of goods or services provided by the data provider" (Draft delegated regulation, 2024, p. 6).

This extended data access via a vetting procedure is particularly interesting for demand side research as it can provide extensive data on platform users and their exposure to political communication. Notably, however, this type of data will only be accessible to researchers working for universities (and publicly funded research institutes) and probably not for independent researchers, NGOs, or (data) journalists.

4.3. European moderation database under DSA 24(5)

To increase the transparency of content moderation digital service providers are mandated to inform the Commission about moderation decisions and the reason for these decisions. All reported moderation decisions are collected in the DSA Transparency Database[7]. While we are not aware of any election research publications that have made use of this resource so far, this can be an interesting data source for investigating and understanding moderation decisions for content related to elections.

4.4. Ad archives under DSA 39

The increased attention to election interference and data protection on digital platforms has previously already resulted in the installation of public advertisement and content libraries (Leerssen et al., 2023). These are intended to open the black box of online campaigning by listing each advertisement's content, sponsor, expenditure, and declared targeting information, giving scholars a reproducible source for large-scale analyses of election advertising (Leerssen et al., 2019). Article 39 of the Digital Services Act obliges very large online platforms to provide such repositories in a searchable format and to supply an API that grants automated access to the same core metadata.

Tools like the Meta and Google ad libraries have existed for some years and enabled researchers to investigate the use of political advertisements on digital platforms. Meta's archive enabled Bär et al. (2024) to study about 80,000 Facebook and Instagram ads from the 2021 German federal election and to detect delivery advantages for some parties. Votta et al. (2024a) mapped 2.5 million Meta ads from 113 elections in 95 countries, showing that micro-targeting is widespread

7. DSA Transparency Database: https://transparency.dsa.ec.europa.eu/

and more sophisticated in highly developed democracies. Kruschinski and Bene (2022) analysed 66,806 Facebook ads from 186 parties in the 2019 European Parliament election and documented strong cross-national variation in paid-media intensity. Extending this work, Kruschinski et al. (2024) combined roughly 10,000 organic posts, sponsored posts, and ads from 53 parties in ten European states and found no systematic differences between populist and mainstream parties in the promotion of divisive topics, negativity, or populist appeals. For the United States, Votta, Noroozian, Dobber, Helberger, and de Vreese (2023) examined the full Facebook and Instagram archive for the seven months before the 2020 presidential vote and reported that official campaigns used more toxic language in narrowly targeted ads, whereas outside groups delivered toxic messages to broader audiences. For Google, Fitzpatrick and von Nostitz (2024) used the ad archive to compare search and YouTube spending during the 2021 German Federal Election.

4.5. Transparency reporting and audits

Digital service providers started reporting about their moderation decisions in voluntary reports before the DSA. However, scope and granularity varied largely between platforms, and currently still vary in the first rounds of risk reports under the DSA. Service providers also report on governments' take down requests and approved content removals.

4.6. Independent data collection beyond APIs and DSA access modalities

As stated before, APIs are not the only option for accessing digital behavioural data. Researchers can also directly cooperate with platforms (or other companies), purchase data (from third-party vendors/resellers), or engage in secondary use of data that has been collected and published by others. If researchers, however, want or have to collect data themselves, the two most prominent alternatives to the use of APIs are web scraping and data donations from users. While the former is a platform-centred approach, the latter is a user-centred one.

4.6.1 Web scraping

Following up on initial suggestions in the papers by Freelon (2018) and Bruns (2019), Mancosu and Vegetti (2020) have advocated for relying more on web scraping for collection platform data in the so-called "post-API age". Despite improvements in data access via APIs brought about by the DSA in the dawning regulated API age, we still believe that web scraping is an important alternative for data access. The importance of web scraping as an independent approach to data collec-

tion resonates with a recent report by the European Digital Media Observatory (EDMO)[8]that has outlined data quality issues of the TikTok API before the 2024 EU elections[9], which they could only test systematically by comparing API data with data collected via a web scraper from TikTok's user interface. Generally, platforms have different policies regarding scraping and, hence, also different mechanisms in place for detecting and preventing or blocking scraping. If or how scraping works depends on the platform and scope of data collection as well as possible other factors, such as location as there may be differences in the way content is presented or is accessible between countries (e.g., with regard to whether a logged-in account is required). To address this issue, again, agreements or cooperations with platforms, and/or legal regulations for the use of web scraping for academic research may be a way forward. A registration of scrapers with exemptions on rate limits could, e.g., be an option here. This, however, would necessitate a process of access control by platforms again. Notably, from a legal perspective it is by now widely accepted that scraping can also be protected under Article 40(12) (Leerssen, 2024). However, for many researchers legal uncertainties and concerns about potential legal actions by service providers remain.

4.6.2. Data donation

A data collection method that has recently become more established in the (computational) social sciences that can also be used for election research are data donations (Hase et al., 2024). Users retrieve their own data from platforms and other digital services based on the General Data Protection Regulation (GDPR) and subsequently share it with researchers via tools such as PORT (Boeschoten et al., 2022). Of course, this approach is also associated with specific challenges and limitations (van Driel et al., 2022), such as sampling biases (see, e.g., Kmetty et al., 2024; Pfiffner & Friemel, 2023; Silber et al., 2022). Similar to data collected via platform APIs, researchers have also found large heterogeneity in the accessibility, content, and quality of the so-called "data download packages" (DDP) that data donation studies make use of. In a recent policy paper, Hase and colleagues (2024) address this issue and present suggestions on how platforms can fulfill data access obligations in a way that facilitates data donation studies. Finally, this approach may be especially difficult to implement for supply-side studies as recruiting the political actors that are investigated as participants and getting them to consent

8. Report on EDMO Workshop on Platform Data Access for Researchers, https://edmo.eu/wp-content/uploads/2024/09/Report-on-EDMO-Workshop-on-Platform-Data-Access-for-Researchers.pdf

9. Researcher Data Access Under the DSA: Lessons from TikTok's API Issues during the 2024 European Elections, https://www.techpolicy.press/-researcher-data-access-under-the-dsa-lessons-from-tiktoks-api-issues-during-the-2024-european-elections/

to and actively engage in data sharing likely poses a serious challenge. However, some work demonstrates that political actors can be willing to share internal campaign data for research purposes. For example, Votta et al. (2025) and Votta et al. (2024b) obtained internal ad data from Meta Ad Manager directly from political parties to examine how ad delivery algorithms affect pricing disparities across parties and audiences during the 2022 Dutch municipal elections and the 2024 European Parliament elections respectively. Or Hewitt et al. (2024) gained access to 617 advertisements from 51 US campaigns through the platform Swayable to analyze ad persuasion effects. These studies remain exceptions, as political parties are often reluctant to share strategic campaign data due to competitive concerns, and establishing the necessary trust requires pre-existing relationships and significant time investment from researchers. Hence, data donation approaches are probably better suited for demand-side studies.

Challenges and promises for election research in the age of regulated data access

Despite the improvements in access to platform data for research brought about by the DSA at the dawn of what we call the regulated API age, researchers face a number of challenges when applying for access and also, after successful application, or when working with data retrieved via APIs. Many of those have been documented in a study by Mimizuka and colleagues (2025) based on interviews with researchers who have been working with online platform data. They also present a set of recommendations for platforms, researchers, and policymakers. The suggestions we present in the following can be seen as complements to those.

The first key challenge for election research in the age of regulated data access is related to the degree of control that platforms have over data access and decision procedures on researchers' applications. For researchers, the decision process can be intransparent and some platforms, such as X, hinder researchers by posing extremely detailed or redundant questions for clarification on data storage or data protection issues (Jaursch et al., 2024; Mimizuka et al., 2025). Considering that researchers' projects and/or funding applications can be conditional on having access to digital platform data, this constitutes a severe challenge. Additionally, access solutions for "non-public data" via data access requests to the national Digital Service Coordinators (DSCs) are not yet in place, which increases researchers' dependency on data access under 40(12) that is controlled by digital platforms themselves. One option for addressing challenges in the data access application process issue could be the involvement of trusted third-parties in data access requests. A

good example in that regard is the collaboration between the Social Media Archive (SOMAR) at the Inter-university Consortium for Political and Social Research (ICPSR) and Meta where SOMAR/ICPSR acts as a neutral and independent third-party for handling requests for access to the Meta Content Library (which is meant to replace the platform CrowdTangle for academic data access). However, also this model may not be optimal as this requires substantial resources and a well-defined process on the side of the third-party. The roughly six-month outage of the ICPSR application platform from Dec 24 to May 25 demonstrates the potential risks associated with this solution. In general, a good approach for addressing issues related to data access applications would be the development of transparent and standardised data application and appeal procedures that are not fully controlled by platforms.

There also are several limitations and challenges related to the use of platform-provided ad archives for research. First, platforms apply inconsistent definitions of what constitutes a political ad, so issue-based or local messages can be omitted. Second, advertiser verification remains weak; declared buyers may mask the real funder. Third, the databases reveal only headline targeting parameters and provide no demographic breakdown of actual delivery, which prevents a full audit of algorithmic biases. Fourth, technical barriers such as rate limits, missing fields, and unstable search interfaces hinder large-scale data collection. TikTok's repository, launched to meet DSA requirements, illustrates these shortcomings: the European Commission found it non-compliant because key data were missing and search was unreliable (European Commission, 2024).

Even bigger issues for research based on ad archives will likely be caused by the decision of both Meta and Google that, from October 2025, they will no longer accept political, electoral or so-called social-issue ads anywhere in the European Union, citing compliance risks under the forthcoming EU Transparency and Targeting of Political Advertising Regulation that supplements the DSA (Meta, 2025; Google, 2024). This retreat reshapes the empirical landscape for online advertising research in different ways: Parties, campaigns and outside groups can be expected to redirect budget to (a) organic posts, (b) smaller or emerging social networks that still permit political ads, and (c) programmatic ad exchanges that fall outside the major-platform ban. With the principal ad libraries drying up, demand will grow for (1) stronger enforcement of Article 39 obligations on every platform, large or small; (2) access to delivery data for programmatic ads; and (3) independent, platform-agnostic monitoring projects (e.g., browser plug-ins or crowd-sourcing schemes) that capture paid messages wherever they appear.

In general, to mitigate the limitations and risks associated with the control that platforms have over data access, researchers need to continue developing and using methods that are less (directly) dependent on the platforms and their decisions (web scraping, data donations) to address the issue of the volatility of the data access situation and the unsuitability of API-based collection for some types of research.

Even if researchers have access to platform data (via APIs), another crucial challenge is testing and ensuring the quality of provided data. For example, researchers have repeatedly identified data quality issues with the TikTok API (Corso et al., 2024; Darius, 2024; Pearson et al., 2024). Data quality cannot be guaranteed if platforms do not have strict reporting and testing guidelines. Importantly, individual research projects often do not have the possibility or capacity to systematically test the quality of data collected by querying research APIs. Coordinated independent data quality audits are a way in which this issue can be addressed.

Besides data access, platforms often also put restrictions on how the data can be used and especially whether and how data sets can be shared with other academics. Principles of open science and also requirements with regard to the reproducibility and replicability of research necessitate that researchers can make platform data available to others. However, such restrictions make adherence to these principles very difficult or even impossible. Twitter/X, for example, has, e.g., only allowed the sharing of post and account IDs. If posts or accounts are deleted, research becomes unreproducible and unreplicable. An analysis by Küpfer (2024) has shown that this issue is especially pronounced for the study of sensitive topics. Accordingly, in a recent opinion piece Davidson et al. (2024) describe platform-controlled social media APIs as a threat to open science. Addressing this issue requires discussing and contributing to the development of new data access solutions, such as remote access/remote code execution (see, e.g., van Atteveldt et al., 2020).

In addition to actively contributing to the development of solutions for the identified challenges, working towards improvements in platform data access in the regulated API age requires that researchers engage in lobbying and collaborate with regulators, policymakers, platforms and relevant institutions. There are positive examples in which policymakers and institutions have been reaching out to the academic community. A prominent one is the consultation on the draft of the delegated act for the DSA. On the national level, the Digital Services Coordinator (DSC) in charge of data access requests under the DSA in Germany (the *Bundesnetzagentur*) has been contacting researchers (including some of the authors of this paper) to

discuss the issue of data access before the German federal election in February 2025. Besides consultations and other forms of direct engagement with policy-makers and institutions, joint policy papers like the one by Klinger and Ohme (2023) and Hase et al. (2024) are relevant initiatives to make the views and needs of researchers visible and clear. Importantly, lobbying and policy work require co-ordination among researchers. Two positive examples in this regard are the activities in this area by DSA 40 Data Access Collaboratory (see Jaursch et al., 2024) and the Coalition for Independent Technology Researchers (CITR).

Overall, we see the emergence of a new age of regulated data access via APIs and other access modalities as a valuable opportunity for election research. As discussed above, there are several remaining challenges, in addition to political resistance from the US. However, we are optimistic that the further implementation of the DSA, coupled with researchers engaging in the activities outlined here (and elsewhere; see, e.g., Mimizuka et al., 2025) can enable novel and high-quality research on elections to better understand both the risks and opportunities of online platforms and their use for political communication.

References

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, *110*(3), 629–676. https://doi.org/10.1257/aer.20190658

Bär, D., Pierri, F., De Francisci Morales, G., & Feuerriegel, S. (2024). Systematic discrepancies in the delivery of political ads on Facebook and Instagram. *PNAS Nexus*, *3*(7). https://doi.org/10.1093/pnasnexus/pgae247

Barberá, P., Casas, A., Nagler, J., Egan, P. J., Bonneau, R., Jost, J. T., & Tucker, J. A. (2019). Who leads? Who follows? Measuring issue attention and agenda setting by legislators and the mass public using social media data. *American Political Science Review*, *113*(4), 883–901. https://doi.org/10.1017/S0003055419000352

Binder, A., Stubenvoll, M., Hirsch, M., & Matthes, J. (2022). Why am I getting this ad? How the degree of targeting disclosures and political fit affect persuasion knowledge, party evaluation, and online privacy behaviors. *Journal of Advertising*, *51*(2), 206–222. https://doi.org/10.1080/00913367.2021.2015727

Boeschoten, L., Ausloos, J., Möller, J. E., Araujo, T., & Oberski, D. L. (2022). A framework for privacy preserving digital trace data collection through data donation. *Computational Communication Research*, *4*(2), 388–423. https://doi.org/10.5117/CCR2022.2.002.BOES

Bouchaud, P., Chavalarias, D., & Panahi, M. (2023). Crowdsourced audit of Twitter's recommender systems. *Nature Science Report*, *13*. https://doi.org/10.1038/s41598-023-43980-4

Bouchaud, P., & Liénard, J. (2024). Beyond the guidelines: Assessing Meta's political ad moderation in the EU. *Proceedings of the 2024 ACM on Internet Measurement Conference (IMC '24)*, 480–487. http

s://doi.org/10.1145/3646547.3689020

Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications*, *10*(7). https://doi.org/10.1038/s41467-018-07761-2

Brautović, M., & Roško, M. (2024). *Generative AI use and disinformation during the Croatian parliament elections 2024*. Adria Digital Media Observatory.

Breuer, J., Bishop, L., & Kinder-Kurlanda, K. (2020). The practical and ethical challenges in acquiring and sharing digital trace data: Negotiating public-private partnerships. *New Media & Society*, *22*(11), 2058–2080. https://doi.org/10.1177/1461444820924622

Breuer, J., Kmetty, Z., Haim, M., & Stier, S. (2023). User-centric approaches for collecting Facebook data in the 'post-API age': Experiences from two studies and recommendations for future research. *Information, Communication & Society*, *26*(14), 2649–2668. https://doi.org/10.1080/1369118X.2022.2097015

Breuer, J., Weller, K., & Kinder-Kurlanda, K. (2023). The Role of participants in online privacy research: Ethical and practical consideration. In *The Routledge Handbook of Privacy and Social Media* (pp. 314–323).

Bruns, A. (2019). After the 'APIcalypse': Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, *22*(11), 1544–1566. https://doi.org/10.1080/1369118X.2019.1637447

Bruns, A., Angus, D., & Graham, T. (2021). Twitter campaigning strategies in Australian federal elections 2013-2019. *Social Media + Society*, *7*(4). https://doi.org/10.1177/20563051211063462

Ceron, A., Curini, L., & Maria Iacus, S. (2016). *Politics and big data: Nowcasting and forecasting elections with social media*. Routledge.

Chu, X., Otto, L., Vliegenthart, R., Lecheler, S., de Vreese, C., & Kruikemeier, S. (2024). On or off topic? Understanding the effects of issue-related political targeted ads. *Information, Communication & Society*, *27*(7), 1378–1404. https://doi.org/10.1080/1369118X.2023.2265978

Corsi, G., Marino, B., & Wong, W. (2024). The spread of synthetic media on X. *Harvard Kennedy School (HKS) Misinformation Review*, *5*(3). https://doi.org/10.37016/mr-2020-140

Corso, F., Pierri, F., & De Francisci Morales, G. (2024). What we can learn from TikTok through its research API. *Companion Publication of the 16th ACM Web Science Conference*, 110–114. https://doi.org/10.1145/3630744.3663611

Darius, P. (2022). Who polarizes Twitter? Ideological polarization, partisan groups and strategic networked campaigning on Twitter during the 2017 and 2021 German Federal elections' Bundestagswahlen'. *Social Network Analysis and Mining*, *12*(1). https://doi.org/10.1007/s13278-022-00958-w

Darius, P. (2024, September 24). *Researcher data access under the DSA: Lessons from TikTok's API issues during the 2024 European elections*. Tech Policy Press.

Darius, P., Drews, W., Neumeier, A., & Riedl, J. (2025). *Radical populist parties receive greater audience support on social media: A cross-platform analysis of digital campaigning for the 2024 European Parliament election*. Center for Open Science. https://doi.org/10.31235/osf.io/42vfx_v1

Darius, P., & Römmele, A. (2023). KI und datengesteuerte Kampagnen: Eine Diskussion der Rolle generativer KI im politischen Wahlkampf. In *Informationsflüsse, Wahlen und Demokratie* (pp.

199–212).

Davidson, B. I., Wischerath, D., Racek, D., Parry, D. A., Godwin, E., Hinds, J., Van Der Linden, D., Roscoe, J. F., Ayravainen, L., & Cork, A. G. (2023). Platform-controlled social media APIs threaten open science. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-023-01750-2

De Vreese, C., & Tromble, R. (2023). The data abyss: How lack of data access leaves research and society in the dark. *Political Communication*, *40*(3), 356–360. https://doi.org/10.1080/10584609.2023.2207488

Digital Services Act, Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and Amending Directive 2000/31/EC (Digital Services Act) (2022).

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

Dommett, K., Power, S., Barclay, A., & Macintyre, A. (2025). Understanding the modern election campaign: Analysing campaign eras through financial transparency disclosures at the 2019 UK general election. *Government and Opposition*, *60*(1), 141–167. https://doi.org/10.1017/gov.2024.3

Eady, G., Paskhalis, T., Zilinsky, J., Bonneau, R., Nagler, J., & Tucker, J. A. (2023). Exposure to the Russian Internet Research Agency foreign influence campaign on Twitter in the 2016 US election and its relationship to attitudes and voting behavior. *Nature Communication*, *14*(62). https://doi.org/10.1038/s41467-022-35576-9

European Commission. (2024, February 19). *Commission opens formal proceedings against TikTok under the Digital Services Act*.

European Parliament and Council of European Union. (2024). *Draft delegated regulation—Ares(2024)7652659 supplementing Regulation (EU) 2022/2065 of the European Parliament and of the Council by laying down the technical conditions and procedures under which providers of very large online platforms and of very large online search engines are to share data pursuant to Article 40 of Regulation (EU) 2022/2065*.

Fitzpatrick, J., & Von Nostitz, F.-C. (2024). Reaching the voters: Parties' use of Google ads in the 2021 German federal election. *Media and Communication*, *12*. https://doi.org/10.17645/mac.8543

Freelon, D. (2018). Computational research in the post-API age. *Political Communication*, *35*(4), 665–668. https://doi.org/10.1080/10584609.2018.1477506

Ghosh, A., Venkatadri, G., & Mislove, A. (2019). Analyzing political advertisers' use of Face-book's targeting features. *IEEE Workshop on Technology and Consumer Protection (ConPro'19)*.

Gibson, R., Bon, E., Darius, P., & Smyth, P. (2023). Are online political influencers accelerating Ddemocratic deconsolidation? *Media and Communication*, *11*(3), 175–186. https://doi.org/10.17645/mac.v11i3.6813

Gibson, R. K. (2015). Party change, social media and the rise of 'citizen-initiated' campaigning. *Party Politics*, *21*(2), 183–197. https://doi.org/10.1177/1354068812472575

Gilardi, F., Gessler, T., Kubli, M., & Müller, S. (2022). Social media and political agenda setting. *Political Communication*, *39*(1), 39–60. https://doi.org/10.1080/10584609.2021.1910390

Global Witness Briefing. (2022, October 21). *TikTok and Facebook fail to detect election disinformation in the US, while YouTube succeeds*. Global Witness.

González-Bailón, S., Lazer, D., Barberá, P., Zhang, M., Allcott, H., Brown, T., & Tucker, J. A. (2023). Asymmetric ideological segregation in exposure to political news on Facebook. *Science*, *381*(6656), 392–398. https://doi.org/10.1126/science.ade7138

Google. (2024, November 14). *An update on political advertising in the European Union*. Google Press Release. https://blog.google/around-the-globe/google-europe/political-advertising-in-eu/

Greene, K. T. (2024). Partisan differences in the sharing of low-quality news sources by US political elites. *Political Communication*, *41*(3), 373–392. https://doi.org/10.1080/10584609.2024.2306214

Guess, A. M. (2021). (Almost) everything in moderation: New evidence on Americans' online media diets. *American Journal of Political Science*, *65*(4), 1007–1022. https://doi.org/10.1111/ajps.12589

Guess, A. M., Malhotra, N., Pan, J., Barberá, P., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Kennedy, E., Kim, Y. M., Lazer, D., Moehler, D., Nyhan, B., Rivera, C. V., Settle, J., Thomas, D. R., & Tucker, J. A. (2023). How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*, *381*(6656), 398–404. https://doi.org/10.1126/science.abp9364

Halavais, A. (2019). Overcoming terms of service: A proposal for ethical distributed research. *Information, Communication & Society*, *22*(11), 1567–1581. https://doi.org/10.1080/1369118X.2019.1627386

Hase, V., Ausloos, J., Boeschoten, L., Pfiffner, N., Janssen, H., Araujo, T., Carrière, T., de Vreese, C., Haßler, J., Loecherbach, F., Kmetty, Z., Möller, J., Ohme, J., Schmidbauer, E., Struminskaya, B., Trilling, D., Welbers, K., & Haim, M. (2024). Fulfilling data access obligations: How could (and should) platforms facilitate data donation studies? *Internet Policy Review*, *13*(3). https://doi.org/10.14763/2024.3.1793

Haßler, J., Magin, M., Russmann, U., & Fenoll, V. (2021). *Campaigning on Facebook in the 2019 European parliament election: Informing, interacting with, and mobilising voters*. Palgrave Macmillan.

Haßler, J., Wurst, A.-K., Pohl, K., & Kruschinski, S. (2024). A consistent picture? Issue-based campaigning on Facebook during the 2021 German federal election campaign. *Politics and Governance*, *12*. https://doi.org/10.17645/pag.8150

Hewitt, L., Broockman, D., Coppock, A., Tappin, B. M., Slezak, J., Coffman, V., & Hamidian, M. (2024). How experiments help campaigns persuade voters: Evidence from a large archive of campaigns' own experiments. *American Political Science Review*, *118*(4), 2021–2039. https://doi.org/10.1017/S0003055423001387

Hjorth, F., & Adler-Nissen, R. (2019). Ideological asymmetry in the reach of Pro Russian digital disinformation to United States audiences. *Journal of Communication*, *69*(2), 168–192. https://doi.org/10.1093/joc/jqz006

Huszár, F., Ktena, S. I., O'Brien, C., Belli, L., Schlaikjer, A., & Hardt, M. (2022). Algorithmic amplification of politics on Twitter. *Procedings of the National Academy of Sciences*, *119*(1). https://doi.org/10.1073/pnas.2025334119

Jaursch, J., Ohme, J., & Klinger, U. (2024). *Enabling research with publicly accessible platform data: Early DSA compliance issues and suggestions for improvement*. Weizenbaum Institute. https://doi.org/10.34669/WI.WPP/9

Jungherr, A. (2016). Four functions of digital tools in election campaigns: The German case. *The International Journal of Press/Politics*, *21*(3), 358–377. https://doi.org/10.1177/1940161216642597

Jungherr, A., Rivero, G., & Gayo-Avello, D. (2020). *Retooling politics: How digital media are shaping democracy*. Cambridge University Press.

Jungherr, A., Schroeder, R., & Stier, S. (2019). Digital media and the surge of political outsiders: Explaining the success of political challengers in the United States, Germany, and China. *Social Media + Society*, *5*(3). https://doi.org/10.1177/2056305119875439

Klinger, U., & Ohme, J. (2023). *What the scientific community needs from data access under Art. 40 DSA: 20 Points on infrastructures, participation, transparency, and funding*. https://doi.org/10.34669/WI.WPP/8.2

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*(15), 5802–5805. https://doi.org/10.1073/pnas.1218772110

Kruschinski, S., & Bene, M. (2022). In Varietate Concordia?! Political parties' digital political marketing on Facebook across 28 countries in the 2019 European election campaign. *European Union Politics*, *23*(1), 43–65. https://doi.org/10.1177/14651165211040728

Kruschinski, S., Bene, M., Haßler, J., Rußmann, U., Lilleker, D., Balaban, D. C., Baranowski, P., Ceron, A., Fenoll, V., & Jackson, D. (2024). Divisive, negative, and populist?! An empirical analysis of European populist and mainstream parties' use of digital political advertisements on Facebook. *International Journal of Communication*, *18*, 5518–5539.

Kruschinski, S., Jost, P., Fecher, H., & Scherer, T. (2025). *Künstliche Intelligenz in politischen Kampagnen. Akzeptanz, Wahrnehmung und Wirkung*. Otto Brenner Stiftung.

Kruschinski, S., & Votta, F. (2025, November 9). *CampAIgnTracker.de: Detection and analysis of visual GenAI in the 2025 German federal election*. https://doi.org/10.17605/OSF.IO/CZEB6

Küpfer, A. (2024). NonRandom Tweet mortality and data access restrictions: Compromising the replication of sensitive Twitter studies. *Political Analysis*, 1–14. https://doi.org/10.1017/pan.2024.7

Lasser, J., Aroyehun, S. T., Carrella, F., Simchon, A., Garcia, D., & Lewandowsky, S. (2023). From alternative conceptions of honesty to alternative facts in communications by US politicians. *Nature Human Behaviour*, *7*(12), 2140–2151. https://doi.org/10.1038/s41562-023-01691-w

Leerssen, P. (2024). *Outside the Black Box: From algorithmic transparency to platform observability in the Digital Services Act*. https://doi.org/10.34669/WI.WJDS/4.2.3

Leerssen, P., Dobber, T., Helberger, N., & de Vreese, C. (2023). News from the ad archive: How journalists use the Facebook Ad Library to hold online advertising accountable. *Information, Communication & Society*, *26*(7), 1381–1400. https://doi.org/10.1080/1369118X.2021.2009002

Lehmann, P., Franzmann, S., Al-Gaddooa, D., Burst, T., Ivanusch, C., Regel, S., Riethmüller, F., Volkens, A., Weßels, B., & Zehnter, L. (2024). *The Manifesto data collection. Manifesto Project (MRG/CMP/MARPOR). Version 2024a*. WZB / IfDem. https://doi.org/10.25522/manifesto.mpds.2024a

Lewandowsky, S., Robertson, R. E., & DiResta, R. (2023). Challenges in understanding human-algorithm entanglement during online information consumption. *Perspectives on Psychological Science*, *19*(5), 758–766. https://doi.org/10.1177/17456916231180809

Maier, J., Stier, S., & Oschatz, C. (2023). Are candidates rational when it comes to negative campaigning? Empirical evidence from three German candidate surveys. *Party Politics*, *29*(4), 766–779. https://doi.org/10.1177/13540688221085239

Mancosu, M., & Vegetti, F. (2020). What you can scrape and what is right to scrape: A proposal for a tool to collect public Facebook data. *Social Media + Society*, *6*(3). https://doi.org/10.1177/2056305120940703

Mangold, F., Schoch, D., & Stier, S. (2024). Ideological self-selection in online news exposure: Evidence from Europe and the US. *Science Advances*, *10*(37). https://doi.org/10.1126/sciadv.adg9287

Meta. (2025, July 25). *Ending political, electoral and social issue advertising in the EU in response to incoming European regulation*. Meta Press Release. https://about.fb.com/news/2025/07/ending-political-electoral-and-social-issue-advertising-in-the-eu/

Mimizuka, K., Brown, M. A., Yang, K.-C., & Lukito, J. (2025). *Post-Post-API age: Studying digital platforms in scant data access times*. arXiv. https://doi.org/10.48550/arXiv.2505.09877

Mols, F., & Jetten, J. (2020). Understanding support for populist radical right parties: Toward a model that captures both demand-and supply-side factors. *Frontiers in Communication*, *5*. https://doi.org/10.3389/fcomm.2020.557561

Müller, A., Riedl, J., & Drews, W. (2022). Real-time stance detection and issue analysis of the 2021 German federal election campaign on Twitter. In *Electronic Government. EGOV 2022. Lecture Notes in Computer Science, vol 13391*. https://doi.org/10.1007/978-3-031-15086-9_9

Munger, K., Egan, P. J., Nagler, J., Ronen, J., & Tucker, J. (2022). Political knowledge and misinformation in the era of social media: Evidence from the 2015 UK election. *British Journal of Political Science*, *52*(1), 107–127. https://doi.org/10.1017/S0007123420000198

Nai, A. (2020). Going negative, worldwide: Towards a general understanding of determinants and targets of negative campaigning. *Government & Opposition*, *55*(3), 430–455. https://doi.org/10.1017/gov.2018.32

Norris, P. (2000). *A virtuous circle: Political communications in postindustrial societies*. Cambridge University Press.

Pearson, G. D., Silver, N. A., Robinson, J. Y., Azadi, M., Schillo, B. A., & Kreslake, J. M. (2024). Beyond the margin of error: A systematic and replicable audit of the TikTok research API. *Information, Communication & Society*, 1–19.

Pfiffner, N., & Friemel, T. N. (2023). Leveraging data donations for communication research: Exploring drivers behind the willingness to donate. *Communication Methods and Measures*, *17*(3), 227–249. https://doi.org/10.1080/19312458.2023.2176474

Popa, S. A., Fazekas, Z., Braun, D., & Leidecker-Sandmann, M.-M. (2020). Informing the public: How party communication builds opportunity structures. *Political Communication*, *37*(3), 329–349. https://doi.org/10.1080/10584609.2019.1666942

Puschmann, C. (2019). An end to the wild west of social media research: A response to Axel Bruns. *Information, Communication & Society*, *22*(11), 1582–1589. https://doi.org/10.1080/1369118X.2019.1646300

Rauh, C., & Schwalbach, J. (2020). *The ParlSpeech V2 data set: Full-text corpora of 6.3 million parliamentary speeches in the key legislative chambers of nine representative democracies*. Harvard Dataverse. https://doi.org/10.7910/DVN/L4OAKN

Rieder, B. (2009, August 13). *Some Yahoo APIs close, mashups too*.

Rieder, B. (2013). Studying Facebook via data extraction: The Netvizz application. *Proceedings of the*

5th Annual ACM Web Science Conference on - WebSci '13, 346–355. https://doi.org/10.1145/2464464.2464475

Rieder, B. (2018, August 11). *Facebook's App review and how independent research just got a lot harder*.

Riedl, J., Drews, W., & Richter, F. (2024). Avoiding the elephant in the room: Echo chambers and the (de-)politicization of COVID-19 during the 2021 German federal election on Twitter. *Frontiers of Political Science*, *6*. https://doi.org/10.3389/fpos.2024.1509981

Sältzer, M. (2022). Finding the bird's wings: Dimensions of factional conflict on Twitter. *Party Politics*, *28*(1), 61–70. https://doi.org/10.1177/1354068820957960

Schmitt-Beck, R. (2000). *Politische Kommunikation und Wählerverhalten*. VS Verlag für Sozialwissenschaften.

Silber, H., Breuer, J., Beuthner, C., Gummer, T., Keusch, F., Siegers, P., Stier, S., & Weiß, B. (2022). Linking surveys and digital trace data: Insights from two studies on determinants of data sharing behaviour. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *185*, 387–407. https://doi.org/10.1111/rssa.12954

Stier, S., Bleier, A., Lietz, H., & Strohmaier, M. (2018). Election campaigning on social media: Politicians, audiences and the mediation of political communication on Facebook and Twitter. *Political Communication*, *35*(1), 50–74. https://doi.org/10.1177/1461444817709282

Stier, S., Siegers, P., & Breuer, J. (2025). Radical right populism and the media: Evidence from the supply side and demand side of political information in Germany. *European Sociological Review*. https://doi.org/10.1093/esr/jcae051

Törnberg, P., & Chueri, J. (2025). When do parties lie? Misinformation and radical-right populism across 26 countries. *The International Journal of Press/Politics*. https://doi.org/10.1177/19401612241311886

Tromble, R. (2021). Where have all the data gone? A critical reflection on academic digital research in the post-API age. *Social Media + Society*, *7*(1). https://doi.org/10.1177/2056305121988929

Utz, C., Amft, S., Degeling, M., Holz, T., Fahl, S., & Schaub, F. (n.d.). *Privacy rarely considered: Exploring considerations in the adoption of third-party services by websites*. arXiv. https://doi.org/10.48550/arXiv.2203.11387

Van Aelst, P., Strömbäck, J., Aalberg, T., Esser, F., De Vreese, C., Matthes, J., Hopmann, D., Salgado, S., Hubé, N., Stępińska, A., Papathanassopoulos, S., Berganza, R., Legnante, G., Reinemann, C., Sheafer, T., & Stanyer, J. (2017). Political communication in a high-choice media environment: A challenge for democracy? *Annals of the International Communication Association*, *41*(1), 3–27. https://doi.org/10.1080/23808985.2017.1288551

Van Atteveldt, W., Althaus, S., & Wessler, H. (2020). The trouble with sharing your privates: Pursuing ethical open science and collaborative research across national jurisdictions using sensitive data. *Political Communication*, 1–7.

Van Driel, I. I., Giachanou, A., Pouwels, J. L., Boeschoten, L., Beyens, I., & Valkenburg, P. M. (2022). Promises and pitfalls of social media data donations. *Communication Methods and Measures*, 1–17.

Ventura, T., Majumdar, R., Nagler, J., & Tucker, J. (2025). Misinformation beyond traditional feeds: Evidence from a WhatsApp deactivation experiment in Brazil. *The Journal of Politics*. https://doi.org/10.1086/737172

Venturini, T., & Rogers, R. (2019). 'API-based research' or how can digital dociology and journalism studies learn from the Facebook and Cambridge analytica data breach. *Digital Journalism*, *7*(4), 532–540. https://doi.org/10.1080/21670811.2019.1591927

Votta, F., Dobber, T., Guinaudeau, B., Helberger, N., & de Vreese, C. (2025). The cost of reach: Testing the role of ad delivery algorithms in online political campaigns. *Political Communication*, *42*(3), 476–508. https://doi.org/10.1080/10584609.2024.2439317

Votta, F., Kruschinski, S., Dobber, T., Cerroni, A., Sandberg, L., Hove, M. F., & Bene, M. (2024). *Examining ad delivery algorithms in the 2024 EP elections*. https://doi.org/10.17605/OSF.IO/RZQM6

Votta, F., Kruschinski, S., Hove, M., Helberger, N., de Vreese, C., & Dobber, T. (2024). Who does(n't) target you? Mapping the worldwide usage of online political microtargeting. *Journal of Quantitative Description: Digital Media*, *4*. https://doi.org/10.51685/jqd.2024.010

Votta, F., Noroozian, A., Dobber, T., Helberger, N., & de Vreese, C. (2023). Going micro to go negative: Targeting toxicity using Facebook and Instagram ads. *Computational Communication Research*, *5*(1), 1–50. https://doi.org/10.5117/CCR2023.1.001.VOTT

Wagner, C., Stier, S., & Zens, M. (2025). *GESIS guides to digital behavioral data 1*. GESIS – Leibniz Institute for the Social Sciences. https://doi.org/10.60762/ggdbd25001.1.0

Wojcieszak, M., Casas, A., Yu, X., Nagler, J., & Tucker, J. A. (2022). Most users do not follow political elites on Twitter; those who do show overwhelming preferences for ideological congruity. *Science Advances*, *8*(39).

Xenos, M. A., Macafee, T., & Pole, A. (2017). Understanding variations in user response to social media campaigns: A study of Facebook posts in the 2010 US elections. *New Media & Society*, *19*(6), 826–842. https://doi.org/10.1177/1461444815616617

Zarouali, B., Dobber, T., De Pauw, G., & de Vreese, C. (2022). Using a personality-profiling algorithm to investigate political microtargeting: Assessing the persuasion effects of personality-tailored ads on social media. *Communication Research*, *49*(8), 1066–1091. https://doi.org/10.1177/00936502209 61965

Zimmermann, F., & Kohring, M. (2020). Mistrust, disinforming news, and vote choice: A panel survey on the origins and consequences of believing disinformation in the 2017 German parliamentary election. *Political Communication*, *37*(8), 215–237.

Published by

Alexander von Humboldt Institute for Internet and Society

CAIS Research for the Digital Age

in cooperation with

CREATe

cnrs centre internet et société

UOC R&I IN3 Internet interdisciplinary Institute Universitat Oberta de Catalunya

University of Tartu Johan Skytte Institute of Political Studies