ARTICLE

More than news! Mapping the deliberative potential of a political online ecosystem with digital trace data

Lisa Oswald [1✉]

Does the internet facilitate everyday public deliberation? Previous research on this question has largely focused on specific aspects, such as online news media diets or political discussions on social media. However, increasingly complex media environments are composed of different arenas with different respective potential for democracy. While previous work extensively dealt with the *quality* of political discussion online, it is a necessary but overlooked step, to consider the upstream features of digital infrastructure and usage. Using digital trace data from Germany, this study maps out which websites are relevant for online public discourse, introduces a measure of deliberative *potential* along six dimensions (information, communication, participation, connectivity, inclusivity and heterogeneity), and explores different types of websites alongside high level usage patterns. Besides a class of mainstream informational hubs, a class of quality information providers that includes most established public broadcasting sites was found. A third class of niche online forums hosts political discussions among more tightly-knit online communities, supporting previous findings of incidental exposure to political content online. While the mainstream information hubs in the sample attract a much larger volume of clicks, users spend relatively more time consuming political information on quality information sites as well as on niche online forums to engage with politics online. This project takes a more holistic perspective of the diverse ecosystem of online deliberation, while presenting a first quantitative exploration of a deliberative system.

[1] Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany. ✉email: oswald@mpib-berlin.mpg.de

Introduction

The question as to what extent the internet enables (or hinders) public deliberation is a much disputed issue that has, so far, only partially been addressed and from rather specific angles. Researchers with a focus on social media platforms have considered active user communication by analyzing online discussion threads (e.g., Esau et al. 2021; Halpern and Gibbs 2013), whereas researchers with an interest in online news media diets, for example, have examined web browsing histories with a distinct news media frame (Guess 2021).

This study takes a step back and focuses on the concept of deliberative *potential*, examining the infrastructural affordances and audiences of different politically relevant websites in the German political online ecosystem. In order to allow claims about the well researched deliberative *quality* and the substance of discussions, a much more indepth analysis of the communicative acts would be necessary and is nothing this paper speak to (e.g., Steenbergen et al. 2003; Esau et al. 2021). The infrastructural elements of a website, such as the provision of political information, comment sections, petitions, etc. as well as the empirical mapping of political usage are features necessary to examine upstream. For example, a discrepancy between the deliberative potential given affordances and usage and the actual quality of discussions may eventually indicate a form of unused potential of online environments for public deliberation – directly implying the next question of why this is.

In other words, an exclusive focus on deliberative quality and relatedly, toxicity in online public discourse, overlooks important selection effects resulting in skewed participation that is central to public discourse in online environments (Kim et al. 2021). While the communicative acts available for analysis on social media or the comment sections of news outlets are predominantly produced by a highly active minority of users, the majority of those reading along while also forming political opinion remains invisible to the researcher and the public (Bright et al. 2019). An exclusive focus on digital traces of communication also risks overemphasizing policies to limit the impact of a skewed highly active minority while overlooking the unused potential of the silent majority for public discourse online. For a comprehensive understanding of the structural transformation of the public sphere in the digital age, one must go beyond the apparent usage behavior of few, but consider the largely invisible behavior of the majority of the public (Habermas 2021). Instead, the examination of web browsing histories across the broader public offers new perspectives to address these methodological challenges.

Furthermore, the academic discourse concerning the extent of homophily and selective exposure in so-called online 'echo chambers' considerably diverges between disciplines and methodological approaches (Pariser 2011; Sunstein 2002). Studies examining data within one specific platform find robust evidence for homophily in social networks (Cinelli et al., 2021; Cota et al. 2019; Guerrero-Solé and Lopez-Gonzalez 2019; Koiranen et al. 2019; Rivero 2019). However, studies considering data across multiple platforms and media types find evidence of a diversity of exposure (Fletcher and Nielsen 2018; Guess 2021; Lelkes 2020; Strauss et al. 2020; Yang et al. 2020). Facing this dispute, a more holistic and data-driven systematic consideration of online arenas for public discourse can help avoid the underestimation of exposure while, at the same time, avoiding the overestimation of siloed information. In line with theorists of deliberative democracy (e.g., Bächtiger and Parkinson 2019), this study demonstrates that the political usage of the internet or the engagement with political topics online goes far beyond the categories of news media consumption and social media discussion but should be measured more holistically, by mapping the diverse ecosystem of online deliberation.

This project links and expands upon existing streams of research on online communication and information, and focuses on the *deliberative potential* of websites as the structural basis for a constructive online public discourse. Bridging those streams of research are a necessary condition for a systematic and systemic assessment of the online public sphere (Bächtiger and Parkinson 2019). The following three research questions are addressed:

1. Which websites hold potential for online public discourse, including political information consumption and discussion online?
2. How is the political online ecosystem structured along infrastructural and usage characteristics?
3. How does the interplay between user demographics and different classes of websites look like?

Using digital trace data from Germany in combination with survey data and manual content coding to characterize a wide range of politically relevant websites, this study empirically tackles various questions of the online public sphere for the first time. With passive web tracking, the data collection is not geared towards one specific platform or website type (e.g. news media), but provides a more complete picture of online behavior, which is crucial for gaining a more holistic and realistic perspective of the online public sphere. The deliberative potential of websites is considered as a latent construct which is in line with an understanding of deliberation as the summative quality of a deliberative system in which different sites fulfill different democratic functions (Bächtiger and Parkinson 2019; Elstub et al. 2019; Esau et al. 2021). Furthermore, a latent classification of websites goes beyond a xylographic distinction between news sites and social media platforms. Using a broad initial scope together with a latent approach, one does not risk overlooking important, potentially overlapping arenas in which political communication is taking place and where people receive their information online.

Overall, this study shows that only a small proportion of online activity (1%) is concerned with politics. To the disappointment of social scientists, the majority of people spend most of their time accessing various forms of entertainment, shopping and work-related URLs. However, the vast majority of users in the sample (1190 out of 1282 individuals) engaged with some political content during the six-month observation period that included the 2017 national elections in Germany. Originally starting with a web tracking dataset of more than 56 million website visits, without setting a predefined exclusion frame for the type of website and using automated approaches together with manual cross validation, the analysis is narrowed down to a set of 69 central domains featuring content on a wide range of political topics relevant to the German public discourse.

Besides a cluster of highly-popular 'mainstream' sites that are visited by a broad range of users to read and discuss political information, a cluster of public broadcasting and journalistic news outlets was found, the 'quality information providers' that cover the highest density of political information. However, they are not characterized by a diverse user base. A third cluster of niche online forums was identified, often dedicated to specific topics or communities, that are mostly neglected in current analyses of the online public sphere. Manual cross checking confirmed that they host in-depth political discussions among sometimes tightly knit online communities such as fan forums. While the mainstream sites in the sample attract a much larger volume of clicks, users spend relatively more time consuming political information via public broadcasting and online news outlets as well as on niche online forums to discuss politics online.

In other words, while the German deliberative system seems to be a rather small fraction of the wider online environment, the

consumption of political content is not as exclusive as the visible discussion patterns of few very active users on social media may imply. While mainstream platforms are most central to the topical flow of political information consumption within the system, the latent structure of deliberative potential highlights the persisting relevance of high quality public broadcasting as the backbone for democratic deliberation in Germany. Niche online forums provide particular potential for interpersonal trust building through prior exchanges about shared a-political interests as potential basis for the deliberation of conflicting political views among citizens.

Deliberation in online environments

Online communication has often been connected to an increase in affective political polarization, the spread of misinformation and the rise of radical counter publics (Bail 2021; Rathje et al. 2021; Bright 2018; Douglas et al. 2017; LorenzSpreen et al. 2021; Vosoughi et al. 2018).

However, in theory, constructive discussions among informed citizens should help to identify the best arguments for complex societal questions and therefore mitigate opinion polarization (Grönlund et al. 2015; Habermas 1984; Ugarriza and Caluwaerts 2014). While more and more deliberation researchers are concentrating their research efforts in the area of online discussions (Strandberg and Grönlund 2018), contradicting evidence is emerging on the quality of online discussions. While this evidence appears negative in many regards (Anderson et al. 2014; Coe et al. 2014; Sunstein 2002; Ziegele et al. 2020), there are positive exceptions, for example when it comes to discussions in the comments section of online newspapers (Manosevitch and Walker 2009; Rowe 2015).

There is a nuanced empirical discourse around the measurement of deliberative quality, including some widely-established coding schemes and the development of novel, more inclusive criteria of deliberative quality (Steenbergen et al. 2003; Graham, 2008, 2012). Additional concepts discussed in the field are for example story-telling, humor, emotions, power, and the role of non-verbal communication for deliberative democracy (Esau et al. 2021; Gerber et al. 2018; Basu, 1999; Coleman and Moss, 2012; Krause, 2008; Follesdal, 2010; Mendonça et al. 2020). However, there is little empirical investigation of the infrastructural foundation of online deliberation—the potential for deliberation supplied online by websites providing political information and discussion spaces.

A promising but today mostly theoretical development in the field are systemic perspectives on deliberation. Deliberative systems theory, that appears particularly applicable in the digital domain, argues that different arenas fulfil different functions for democracy (Bächtiger and Parkinson, 2019; Ercan et al. 2017; Mansbridge et al. 2012). However, the empirical conceptualization of the deliberative *potential* of websites as the basis for a constructive public discourse remains largely unresolved. Even though online political deliberation might be a niche phenomenon rather than mainstream behavior, it is crucial to understand its structural foundation. Beauchamp (2020) describes the deliberativeness of discussions in online environments as a function of membership and structure. This project empirically maps these structures, an ecosystem of politically relevant websites, as the foundation of a deliberative system and the necessary condition for deliberation to occur in online environments.

Deliberative potential of websites

While the theoretical term 'deliberative potential' is not a novel concept in the field, the deliberative potential of websites has, so far, only been explored theoretically or with regard to specific domains (Conover et al. 2002; Mendonça and Ercan 2015). For example, Wiklund (2005) analyzed different municipal websites in Sweden with a focus on two dimensions, information services and communication services provided by the websites. In contrast, Richardson and Stanyer (2011) examined British online news outlets. They consider manifest characteristics of websites while still keeping a focus on communicative features such as online forums and the deliberative quality of communication.

In this study, the assessment of deliberative potential is rooted in the theory of deliberative democracy; the six dimensions of the concept are described in detail below. The notion that 'different types of public deliberation online can be expected to display different characteristics and fulfill different functions in democratic opinion and will formation, as well as in decision making.' (Esau et al. 2021, p. 2) has especially influenced the definition of deliberative potential used throughout this project. While different types of deliberation, ranging from intimate personal discussion to anonymous public communication fulfil different functions, they are also likely to occur in entirely different arenas that come with distinct infrastructural setups and user constellations. In turn, these arenas are not just the playing field for political discussion but shape discussions with their respective potential (Beauchamp 2020).

The dimensions of deliberative potential are structured along three core functional dimensions that are determined by the infrastructure of the website as the basis for deliberation (see Table 1). Three additional higher-level dimensions are defined by the respective usage patterns or demand-side characteristics. For example, a forum that enables reciprocity in communication is regarded as necessary basis for online deliberation. However, in line with theories of deliberative democracy (Bächtiger and Parkinson 2019; Habermas et al. 1974) only the consideration of heterogeneous arguments within an inclusive debate that is accessible for a diverse citizenship can make a discussion truly deliberative. This conception is not particular to the online sphere. Already in face-to-face citizen assemblies, the deliberative quality as well as the outcomes of deliberation depended on who is deliberating, regardless of the venue's infrastructure (Warren 2021). While from a normative point of view, the combination of certain characteristics is favored, the systemic notion of deliberation does not require all arenas to fulfill all characteristics at the same time (Bächtiger and Parkinson 2019).

Therefore, in this project maps different structural preconditions for deliberative discourse, including both infrastructural aspects and patterns of how this infrastructure is used. In doing so, this project focuses on the description of the deliberative *potential* of online environments, rather than assessing the *quality* or *issues* of the discourse.

Information. The first dimension in the assessment of the deliberative potential of websites is the provision of relevant information. In 1789, Thomas Jefferson had already concluded that well informed citizens are the key to a healthy democracy (Jefferson 1789), a notion that still constitutes a core feature of deliberative democracy (Bächtiger and Parkinson 2019; Fishkin 2018). Information on parties, policies, institutions and procedures are the building blocks of political knowledge and are extensively researched concepts in the social sciences for good reason (Carpini and Keeter 1993; Prior 2005). Civic knowledge about institutions and processes can help citizens to better understand their interests as individuals and members of a group, it increases the consistency of views across issues and over time, and it increases trust, political participation and support for democratic values, such as tolerance for the needs of minorities (e.g. Galston 2001). Moving the perspective to the digital era, in

Table 1 Deliberative Potential Criteria.

Dimension	Operational Definition	Criteria	Measurement
Information	Users can find political information on this website. Such sites serve as a resource for opinion and will formation.	1. Information on political issues, actors, and institutions 2. Administrative or local information 3. Primary or (journalistically) curated source of information	Human rating (binary 0/1) for each criterion
Communication	The website provides users with the possibility to express and/or exchange political opinions with other users. Such sites serve as communicative spaces for (interactional) opinion and will formation.	1. Enables commenting / political expression (and potential rating of comments) 2. Enables reciprocity / replies to comments of other users (open replies, not only ratings)	Human rating (binary 0/1) for each criterion
Participation	The website provides users with the possibility of online political participation or organization, implying a potential (direct) impact on political decision-making.	1. Enables contact to political actors 2. Enables political participation (petitions, polls, etc.) 3. Enables political organization (events, groups, etc.)	Human rating (binary 0/1) for each criterion
Connectivity	The website is connected to other relevant websites. This facilitates further research on political issues or the implementation of intentions of political participation.	1. From this website, users move to other relevant websites 2. Users land on this website after visiting other relevant websites	Network centrality, in/outgoing relevant traffic
Inclusivity	The website is used by a comparably diverse set of individuals. This dimension serves as an indicator for low barriers of access.	1. Education 2. Gender 3. Age	Quantitative assessment using webtracking data in combination with survey information
Heterogeneity	The website connects people holding a comparably wide range of political opinions. This stands in contrast to the notion of ideological online 'echo chambers'.	1. Political orientation 2. Party preferences	Quantitative assessment using webtracking data in combination with survey information

Note. Curation represents a basic form of fact checking taking place (no obvious misinformation or strongly opinionated statements presented as factual statements). Connectivity, inclusivity and heterogeneity are optional categories, as only possible to assess with digital trace data + survey information. Criteria developed with reference to Wiklund (2005); Esau et al. (2021); Richardson and Stanyer (2011); Mansbridge et al. (2012).

the context of online deliberation, websites on which users find political information can serve as a resource for opinion and will formation (Esau et al. 2021). While information on political parties and issues potentially influence political opinions and inform vote choices, administrative information help citizens to understand democratic structures and procedures (Döring 2021).

Drawing upon both empirical findings on political and civic knowledge (Carpini and Keeter 1993; Munzert and Selb 2017) as well as previous research on the role of political information for deliberation (Wiklund 2005), three distinct criteria are included for the assessment to what extent a site provides relevant political information. This study assesses whether the site provides (1) information on political actors, institutions and political issues, (2) information on administrative procedures and local information, and (3) whether information provided by the site is journalistically curated or has, at least, undergone some other form of fact checking (such as e.g. on Wikipedia). Local information, for example on local initiatives and regulations is integrated into the category with administrative information, because they have the similarly enabling potential for civic engagement and political participation. The third information criterion serves as a basic manifest proxy for information quality.

Measuring the exposure to diverse news media is one important component to assess the informative potential of the online sphere for a functioning democracy. Previous projects focused on news access through social media sites which, however, risks neglecting less mainstream media outlets (Bakshy et al. 2015; Barberá et al. 2015; Eady et al. 2019). Other studies that collect data from the user perspective risk biased self-reports in surveys (Boxell et al. 2017; Lelkes, 2016). Facing these methodological challenges, web tracking data provide unique insights into real-life media diets. While Guess (2021)

illuminates important aspects of online news media diets, for example, this study aims to capture the overall deliberative potential of the complex online public sphere using an even more inclusive scope.

Communication. It is important to note that this project does not consider the deliberative quality or the content of discussions taking place on a certain website, when looking at the communication dimension of deliberative potential. Instead, this study assesses whether the website provides users with the possibility to express and/or exchange political opinions with other users (Wiklund 2005). Such sites can serve as communicative spaces for interactional opinion and will formation (Esau et al. 2021). The dimension splits into two levels of communication. First, it is measured whether the website enables the expression of political opinions through the provision of comment sections, for example. In a second step, it is assessed whether the site fulfills the deliberative core criterion of potential reciprocity (Steenbergen et al. 2003). Communicative situations can only be characterized by reciprocity, if users have the option to reply to each other. Therefore, websites only fulfill the second criterion if a reference to previous comments is enabled, for example in online forums and on social media platforms, but also in comment sections of online news papers where 'reply' functions are enabled. The sole possibility to up-vote or down-vote comments, or to react to comments through 'likes' is not regarded as reciprocity. Following this approach, this study does not assess actual communication or specific elements such as listening to the arguments of others (Kriplean et al. 2012; Scudder 2020), but it assesses the structural foundation as preconditions for deliberative communication.

Participation. Websites that provide users with the possibility of online political participation can have a more or less direct impact on political decision-making or, at least, serve as a platform for the aggregation of interests (Esau et al. 2021).

It is a disputed issue, whether a link to decision-making is necessary to consider political communication as deliberation. While Thompson (2008) insists on the link to decision-making, the idea of deliberative polls (Fishkin et al. 2018), the Habermasian idea of diffuse communication in the public sphere as well as the deliberative systems approach adopt a broader definition of deliberation. By mapping the deliberative potential of the online ecosystem, this study includes opportunities for online political participation as desirable features of online political ecosystems without making a judgment about the definition of deliberation itself.

To assess the structural foundation of the link to decision-making, three distinct criteria are examined: (1) whether the website hosts petitions and/or opinion polls to collect, aggregate or organize public opinion (Richardson and Stanyer 2011), (2) whether the website enables citizens to get in contact with political actors (Wiklund, 2005), and (3) whether the website enables the political organization of citizens, for example by the formation of political interest groups or events such as discussion forums, demonstrations and other forms of political protest. Wiklund (2005) considered some of these criteria under the framework of the communicative services of a website. However, it might be worth distinguishing between forums for discussions among citizens and communicative acts that can have a more direct influence on political decision-making.

While this study considers the first three dimensions: information, communication and participation, as core dimensions of the deliberative potential of a platform, three additional criteria are assessed: connectivity, inclusiveness and heterogeneity that are defined through usage patterns and user characteristics.

Connectivity. The dimension of connectivity considers whether the website is connected to other politically relevant websites. These connections can, for example, enable further research by citizens on political issues or facilitate the implementation of intention to participate in the political process. Following the conceptualization of deliberative systems, an understanding of the links and flows between different sites is crucial for a systemic understanding of online public discourse (Dryzek 2012; Esau et al. 2021; Fleuß et al. 2018). For example, Fleuß et al. (2018) emphasized the transmissions between different loci as being an important aspect to measure deliberation in a systemic way. They proposed tracking the transmission of topics as they evolve within the system as well as tracking individuals who transmit ideas from one locus to another. While the analysis mainly operates within the arena of informal deliberation, the approach to operationalize connectivity, by tracking users' subsequent visits to different websites featuring the same topics, gets very close to their theoretical idea of formalizing transmissions.

A body of literature outside the field of deliberation research that examines those links and flows between different online sites can be found in the field of inter-media agendasetting research. For example, media researchers have analyzed to what extent certain online publics are able to stimulate mass media publics, others have examined news diffusion processes from a temporal perspective or the Twitter networks of journalists as important nodes between digital and mass media (Messner and Distaso 2008; Wallsten 2007; Buhl et al. 2018; Neuberger et al. 2019).

In this course, digital trace data provides a unique opportunity to explore linkages between websites through the observation of real-life online behavior with network-analytical quantities. At the time of writing, this project is the first to formalize the connectivity of the different arenas of an online deliberative system empirically.

Inclusivity. The dimension of inclusivity appears to be an increasingly prominent aspect in the scientific discourse on deliberation. Mansbridge et al. (2012) describe three overall functions of a deliberative system: (1) an epistemic function to produce appropriately informed preferences and opinions, in this study, considered under the dimension of information, (2) an ethical function that creates respect between experts and citizens – these aspects could, for example, be a consequence of communication among citizens and contact between citizens and political actors, both captured in the dimensions of communication and participation –, and (3) a democratic function. Mansbridge et al. (2012) conceptualize the democratic function as promoting 'an inclusive political process in terms of equality' (p. 12), which implies the inclusion of multiple and plural voices.

This study explicitly considers the demographic variables of gender, age and educational[1] background in the assessment of inclusivity, to answer the question whether a website is used by a diverse set of individuals. This dimension, with a focus on demographic aspects, features of social groups, corresponds to Young's (2002) concept of diverse perspectives for democratic representation. In the digital context, this dimension can further serve as indicator for low barriers of access. The unique combination of digital trace data with high-quality survey data allows a precise measurement of this dimension of deliberative potential.

Heterogeneity. One final important aspect, also implied in the conceptualization of the democratic functions of a deliberative system by Parkinson and Mansbridge (2012) is the inclusion of a variety of interests, concerns and claims. Furthermore, Young (2002) emphasizes the representation of diverse opinions, "any judgements or belief about how things are or ought to be" (p. 135) for a pluralistic democracy. This aspect is considered under the dimension of heterogeneity and assesses whether the website connects people holding diverse political opinions. This concept stands in contrast to the notion of ideological online 'echo chambers' in which users are argued to be mainly surrounded by similar others, holding opinion-reinforcing views (Pariser 2011; Sunstein 2002). In contrast to other researchers, who focused on the heterogeneity of information diets in online environments, this study considers the composition of users that visit a website (Bright et al. 2020; Dubois and Blank 2018; Guess 2021). More specifically, the approach taken in this study captures the heterogeneity of political orientations of users within a website through their explicit indication of political orientation on a left-to-right scale and their expressed party preferences in the context of the German federal election in 2017.

Both dimensions, inclusivity and heterogeneity are rooted in representation literature (Warren 2021). Random sampling would, under ideal experimental conditions with high compliance, ensure inclusivity and heterogeneity. Under natural conditions in online environments, the issues of inclusivity and heterogeneity as basis for discursive representation are more complex. This study considers the concepts of inclusivity and heterogeneity as theoretically distinct, as inclusivity builds on manifest demographic variables of the user whereas heterogeneity is a more latent construct of political attitudes and party preferences that possibly links more closely to political content featured online.

Methodological challenges

Around the beginning of the millennium, Steyaert (2000) had already emphasized the need for analytical tools that enable a systematic empirical analysis of digital democracy. However, most research in the field is still being conducted as explorative case studies, mostly with a focus on the content or the deliberative quality of communicative acts online (Felicetti et al. 2016; Jensen 2003; Jonsson 2015; Pedrini 2014). Also the rooting theorist of deliberative democracy and the concept of the public sphere, Jürgen Habermas, explicitly emphasized the methodological challenge of empirically examining online deliberation in a recent piece on the restructuring of the public sphere in the digital age (Habermas 2021). The conceptualization of deliberation as the emergent property of a system, involving the dynamics of contexts and platform design elements with different functions for democracy, comes with serious questions for empirical research (Esau et al. 2017; Boswell and Corbett 2017; Fleuß et al. 2018; Niemeyer et al. 2015).

Based on the current state of the empirical literature, this study identifies two key method-ological challenges in the analysis of online deliberation. First, given the ever-growing online landscape, it is crucial to know *where* on the web different branches of public discourse are taking place in order to make assumptions about their content and quality. The question as to which websites are used for political issues is not trivial as deliberation in online environments is getting increasingly pluralistic and incidental news exposure occurs regularly (Esau et al. 2017; Janssen and Kies 2005; Feezell 2018; Kim et al. 2013; Tewksbury et al. 2001; Yadamsuren and Erdelez 2010).

Second, most empirical research in the field of online deliberation, especially the assessment of the deliberative quality of communication, is researched on individual (active) behavior. However, most people on the web are passive consumers of content, also called 'lurkers' (Nonnecke and Preece 1999; Sun et al. 2014). This passive majority does not leave obvious digital traces in online forums and comment sections but they certainly do obtain political information from the web that shape their political opinions and actions. In the context of social media platforms, for example, passive users also experience social learning and constantly adapt their understandings of social norms by observing other people's communication, while a highly active minority, also called 'power users', creates the majority of content online (Bright et al. 2019). This distinction between passive and active engagement in online public discourse has, with a slightly different angle, already been reflected in Habermas' notion of a 'two-track model' of deliberation, emphasizing that most political deliberation happens in institutionalized form with the majority of citizens taking a pure spectator role (Habermas 1996).

In consequence, it remains largely unclear how this silent majority uses the web for political issues facing a heavy focus on communication data. It would be important to also examine passive exposure instead, to meaningfully define the boundaries of the public arena online. The question as to what extent websites *enable* public deliberation, under the further consideration of a systemic perspective, is what this study aims to answer with the assessment of the deliberative potential of websites.

Of course, deliberative potential does not directly imply deliberation. Online environments may provide accurate political information but also misinformation, they may enable deliberative discussion but also host toxic exchanges, they may provide platforms for civic engagement but also hostile participation (Freudenthaler and Wessler 2022; Quandt 2018). However, if the fundamental conditions of deliberative potential are not met in the infrastructure and usage of online environments, deliberation itself is impossible.

Methodology

The analysis is based on web tracking data that was collected within a six-month period in Germany, including the national elections in September 2017. The data is linked to rich survey data, including users' demographics, political attitudes and other relevant political variables. This unique combination of two data sources allows the assessment of the deliberative potential of websites enriched by components that only become measurable in the interaction of user characteristics and usage behavior (connectivity and especially, inclusivity and heterogeneity).

Data. The survey data was collected via the German YouGov Pulse panel with survey questions originally fielded to 1500 respondents in five waves. Using a quota-sampling procedure on the basis of the marginals from Best for Planning (2017), the sample mirrors the German online population with respect to gender, age and, to some degree, education. Respondents were asked to install a URL tracking software that uses passive metering technology to record detailed browser histories on an opt-in basis. Tracking could be paused for 15 min and respondents could end their participation at any time. This digital trace data includes more than 56 million website visits to almost 200,000 different domains by 1282 different individuals between July 2017 and December 2017. As this paper draws on data collected in a prior research project, details on the sampling procedure, the deployment of the passive metering software as well as privacy and ethical considerations can be found in part B of the supplementary information.

In a first step, the top 1000 domains were manually classified into categories (see Table C12). Those 1000 domains account for about 83% of website visits. This distribution is characteristic for web browsing data, in which central websites accumulate most activity while the majority of sites are only visited by very few users. The pre-labelled web tracking data was then merged with the survey data to allow for the description of the demographic profile of the sample[2].

Measures. To develop a measurement for the deliberative potential of websites, the set of websites in scope had to be identified first. Considering the context of deliberative democracy, a focus on websites that, in the wider sense, play a role in the German online public discourse or feature political content appeared the most appropriate. Through this empirical approach, the notion of deliberation underlying the concept of deliberative potential is focused on political topics and set apart from everyday conversation or everyday deliberation that may only implicitly regard political issues (e.g., Maia 2017). The approach to be particularly inclusive in the first step sets this project apart from previous research, for example on online news media diets (Guess 2021), that also uses web tracking data but focuses exclusively on news websites. In order to gain a realistic picture of the online public sphere, it is important to consider all possible channels through which political information, communication and participation is enabled, especially because the exposure to political information makes up only a small proportion of users' total online engagement.

Exploiting the fact that the data almost mirror the German online public demographically and include individuals' browser histories for a period of about six months around the federal elections in Germany, websites accessed in a 'political context' were selected with a combination of automated keyword search, matching political keywords to the URL strings of tracked website visits, and manual cross checking by browsing the websites for instances of clearly political content (see Fig. 1). The relevance of these websites for the online public discourse in Germany in 2017
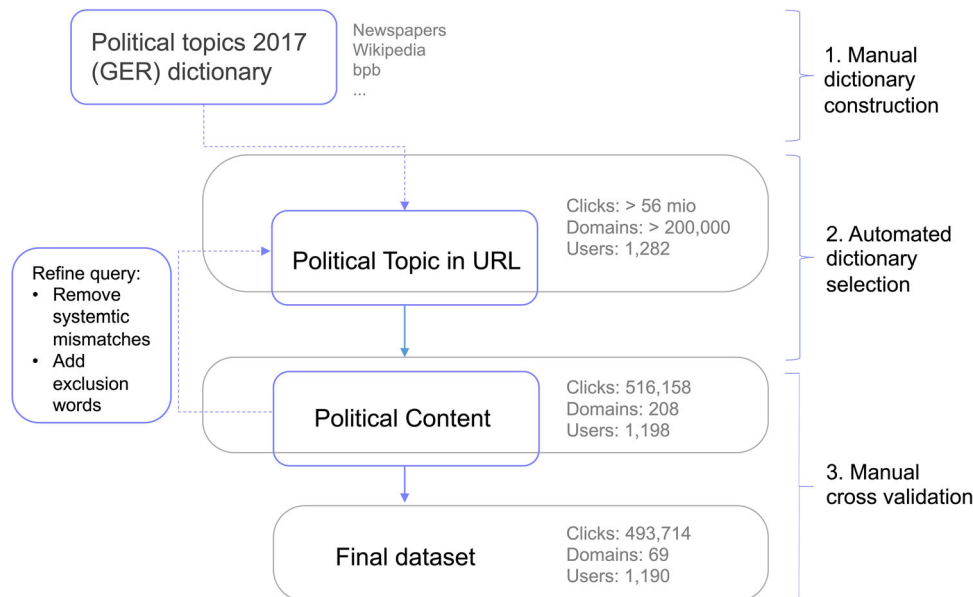
Fig. 1 Topic-driven selection process of websites in scope. Top: Manual construction of dictionary consisting of political topics in the Germany public discourse of 2017. Center: Mapping topic dictionary onto full URLs of website visits as automated selection step. Bottom: Manual cross-validation of automated selection and refinement of dictionary for systematic mismatches.

was then approximated using the number of website visits on the respective domain, aggregated across the sample, as a measure of engagement. The process of political website identification is described in detail in SI A.

In a second step, the deliberative potential of those politically relevant websites was determined. To this end, the outlined six dimensions of deliberative potential were assessed: information, communication, participation, connectivity, inclusivity and heterogeneity (see Table 1). While the first three dimensions were evaluated with manual content coding only, the latter three were determined through a consideration of digital trace data in combination with linked survey data.

Content coding. In order to assess the first three core dimensions of the concept of deliberative potential, the full sample of relevant websites was assessed using quantitative content analysis on the websites' infrastructure. The theoretical definitions of the dimensions were translated into operational definitions including concrete criteria that could be assessed using a binary rating system (0 representing 'not present', 1 representing 'present'). The unit of analysis were website domains and the coding was conducted after manually visiting the website and assessing the overall structure of the page, posts, articles, and comment sections. A standardized code book (see Table 1), including all dimensions and sub-criteria was used to streamline the coding process.

Digital trace data. The availability of web tracking data in combination with survey data allows the enrichment of the manual assessment of deliberative potential with granular quantitative measures of online behavior. This micro-level behavioral data was used to assess criteria on a more macro level, the unit being websites rather than single users. The connectivity measure was constructed through network analytical measures of in-going and out-going traffic (Csardi et al., 2006). The measure of inclusivity was added based on demographic variables; heterogeneity based on the political preferences of users.

More specifically, in order to exploit the benefits of digital trace data to build the connectivity measure, a network of website visits

was constructed with websites represented as nodes, and temporally subsequent website visits for one user, featuring the same topic, represented as edges. For example, if a user reads an online newspaper article featuring the name 'Merkel' in the URL and, following this, visits a social media discussion featuring 'Merkel', an edge was created between the nodes of the online newpaper and the social media platform. Only subsequent visits to different websites were counted, while self-loops were excluded from the network. This way, instead of considering the 'dead' hyperlink-infrastructure from the html text of the websites, a measure of actual in-going and out-going politically-relevant traffic was created for each website. These traffic flows can be quantified using the network analytical measures of in-degrees (in-going traffic) and out-degrees (out-going traffic) (see Fig. C4 in the SI).

To assess the degree of inclusivity of a website, three distinct diversity indicators were calculated for each website for the variables age, gender and education. The widely-used entropy-based Shannon-Wiener diversity index was used as it is implemented in R (see SI A; Dixon 2003; Grafton et al. 2012; Kiernan 2014; Oksanen 2013). A high inclusivity means that a website is accessed by individuals from different age groups, education levels or genders. The more different categories (for example age groups) and the more similar the engagement levels across those different groups, the higher the estimated inclusivity value of a particular website.

For the construction of heterogeneity criteria, a similar approach was used. The diversity assessment was applied to a variable measuring the political orientation of participants on a left-to-right scale and to their reported first votes in the 2017 federal election in Germany. According to the Shannon-Wiener diversity index, the heterogeneity of a website is comparably high if it is visited equally by individuals with different political orientation.

The dataset of individual websites, labeled with regard to the six criteria of deliberative potential, is one outcome of this study which is published along this manuscript. However, this dataset needs to be structured and summarised to be digestible and informative. The reduction of complexity by structuring data is
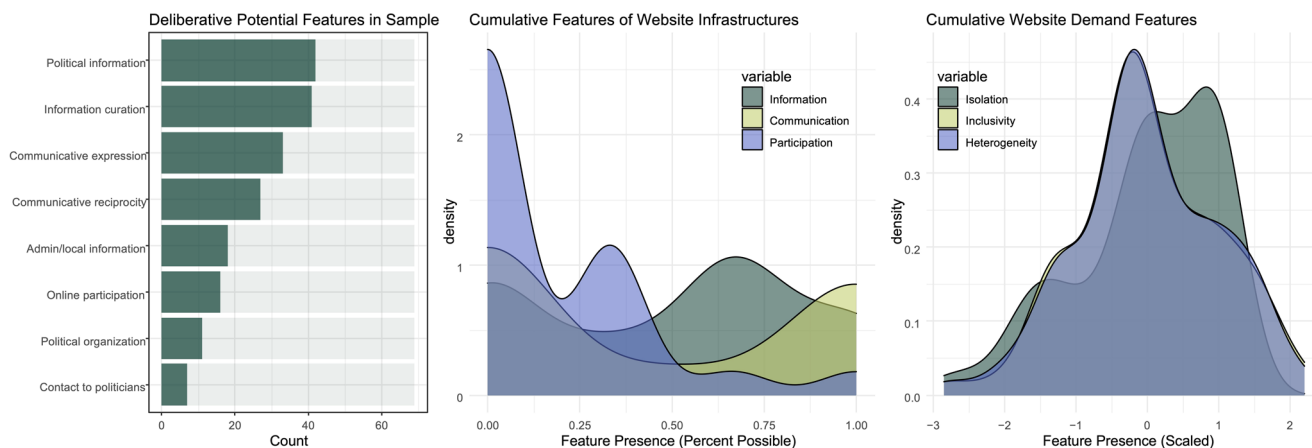
Fig. 2 Deliberative feature descriptives. Left: How many of the 69 websites fulfill criteria? Center: How does the cumulative feature presence look like for the three infrastructural criteria? E.g. most websites fulfill 0 out of 4 participation criteria, 2 out of 3 information criteria and either 2/2 or 0/2 communication criteria. Right: How does the cumulative feature presence look like for the three usage-based criteria? Metrics scaled for better comparability.

the core purpose of clustering approaches, including latent class analysis, which is why it was used in this manuscript in a second step, after the rich classification of each website along six deliberative criteria.

Clustering websites with latent class analysis. After the assessment of all six dimensions of deliberative potential of websites, patterns of commonalities and differences were considered between websites to explore different 'profiles' of deliberative potential. In line with the latent understanding of deliberative potential, a latent class approach was used to identify groups of websites according to their deliberative potential. Besides this theoretical reason, an examination of the empirical relationships between different criteria, suggests the use of a latent composite measure as there are both, correlations within, but also between different dimensions of deliberative potential (see Fig. C5). More details on the latent class modeling approach can be found in SI A. Finally, after the identification of classes, an individual class membership prediction value was assigned to each website, allowing the categorization of websites into latent classes.

Results

Politically relevant sites in Germany in 2017. Applying the two-stage process of website selection, consisting of the automated dictionary-based classification of websites as 'politically relevant' and the following manual cross validation, 69 central domains were identified in the sample that have played a role in the online public discourse in Germany in the second half of 2017. Those websites were visited by 1190 unique users, which included a large proportion of the original sample ($N = 1282$). It is important to note that this does not mean that, for example because highly frequented websites such as 'Google' and 'Facebook' are part of this set of 69 websites, those 1190 individuals simply used those platforms at least once in the six-month period. Instead, it means that they 'googled' some political keyword or visited political content on Facebook because the political filtering step took place *before* the compression of website visits into domains.

Starting from the original sample of more than 56 million tracked website visits, less than 1% (493,714 clicks) were politically relevant visits to those 69 domains. Table C12 summarizes the big picture of the overall online activity of the sample, illustrating that the engagement with political issues is not the dominant motive for many users to use the web. In

contrast, the most frequented websites were social media platforms and search engines (mostly for apolitical content), email providers, online shopping, gaming, streaming, porn and online banking.

Furthermore, only slightly more than half of the politically relevant websites (52%) in the sample are explicitly labeled as news websites, and only 12% of the websites featuring political discussions are social media platforms.

Figure 2 summarizes the descriptives on the prevalence of deliberative potential criteria across the sample of politically relevant websites. While the majority of websites fulfills two out of three information criteria (most provide political information that underlie some form of journalistic curation or fact checking), only very few websites fulfill the criteria of participation. When it comes to the potential to host political discussions, about half of the platforms provide the possibility to express *and* discuss political opinions online while the other half *neither* enables expression nor reciprocity in communication. Only very few platforms enable the expression of political opinions in the form of comment sections without the possibility to reply to other comments. Considering the 'demand side' characteristics of demographic inclusivity and political opinion heterogeneity within websites' user bases, both measured with the entropy-based Shannon-Wiener diversity index (more details see SI A), there is a considerable overlap of density distributions. Websites attracting users of diverse age groups, genders and education levels appear to also attract users of diverse political orientations and party preferences. The distributions of both measures, though highly correlated with the overall engagement on a website, does not mirror the rather leftskewed metric of connectivity that reflects engagement links and flows between politically relevant platforms.

To structure the political online environment along the complex set of deliberative potential criteria, a latent class analysis was conducted. Considering various model fit criteria and rounds of validation, a model with three latent classes was selected (see Fig. C6, Table C2 and more description in the SI).

Latent class structure of the online ecosystem. Figure 3 summarizes the conditional probabilities of websites belonging to each of the three latent classes dependent on their fulfillment of each of the deliberative potential criteria. It also present exemplary sets of websites that were previously identified as politically relevant and sorted into the three estimated latent classes based
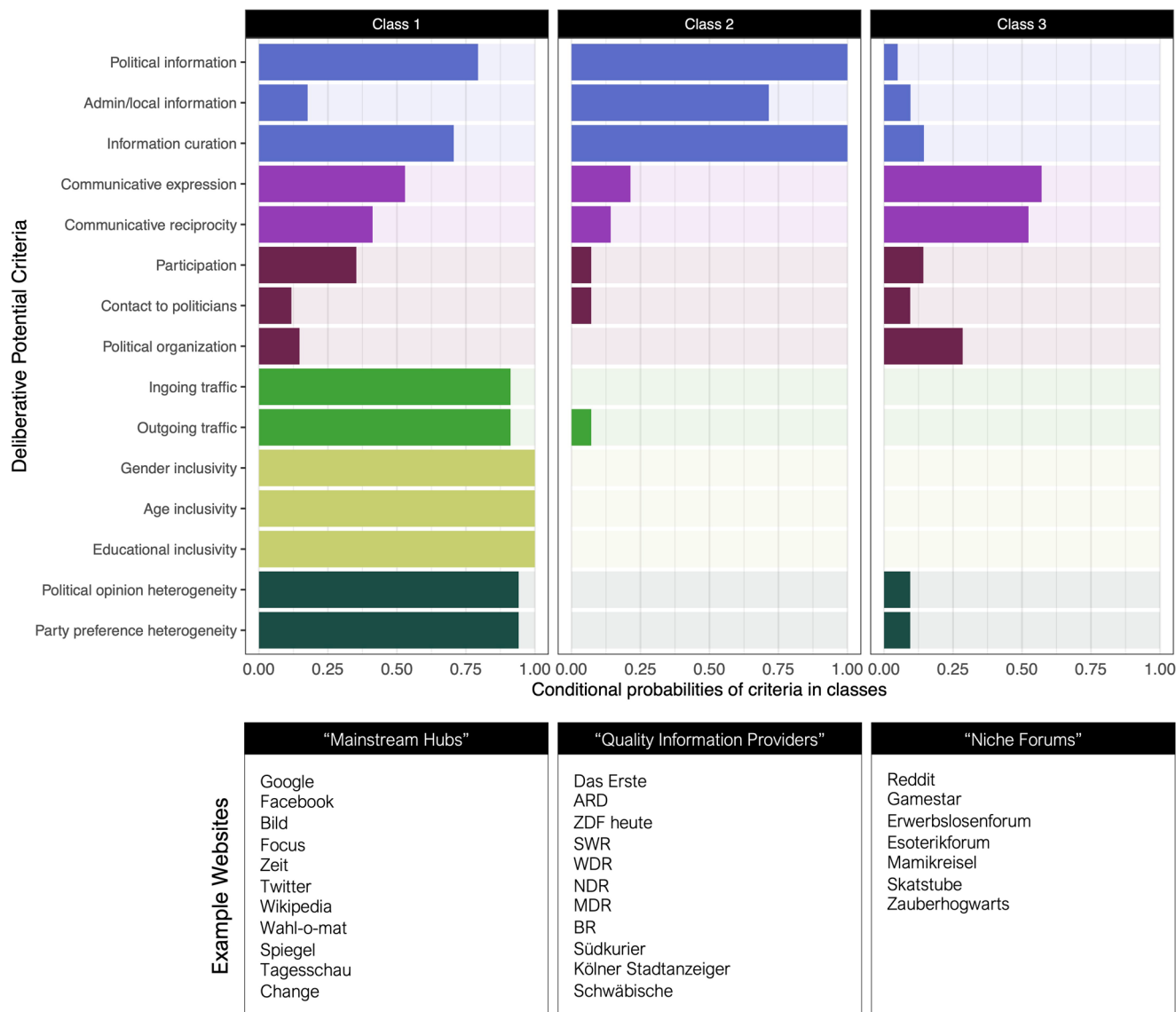
Fig. 3 Conditional response probabilities, by deliberative potential criterion, of belonging to each latent class. Based on response probability patterns and class membership, class 1 was named 'mainstream hubs', class 2 was named 'quality information providers' and class 3 was named 'niche forums'. Full list of domains provided in SI Table C1.

on their respective predicted probabilities of class membership (See SI Table C1 for the full lists).

In summary, websites in class 1, from now on referred to as the 'mainstream hubs', show especially high class-conditional probabilities of fulfilling the dimensions of connectivity, inclusivity and heterogeneity, while websites assigned to class 2, the 'quality information providers' appear strong with regard to information criteria. Websites assigned to class 3, the 'niche forums' show rather low class-conditional probabilities for most criteria of deliberative potential, except for the communication dimension and political organization.

More specifically, the class of *mainstream hubs* (class 1) is composed of a diverse set of websites that fulfill the core criteria of information, communication and participation to some extent but which are especially characterized by a high degree of connectivity, demographic inclusivity and political opinion heterogeneity. Overall, those websites have the highest level of engagement measured by the number of website visits in the sample. Such sites are, for example, prominent high quality

national newspapers like 'Zeit' and 'Spiegel', more tabloid outlets like 'Bild', social media platforms like 'Facebook' or 'Twitter', but also sites with particular functions, such as the online petitioning platform 'Change' or the voting advice application 'Wahl-O-Mat'. What most of the websites in this class have in common is that they are highly-frequented websites that are nationally well known and relevant for political content across diverse German-speaking audiences.

The *quality information providers* (class 2) include almost exclusively established local, regional and national online news outlets and informative TV channels hosted by public service broadcasting with the exception of 'RTL' and 'Sat1', two private TV channels with broad online news sections. While 'ARD' is the leading national public service broadcasting channel in Germany, 'MDR', 'WDR', 'SWR' and 'NDR' are their regional channels. Websites like 'Südkurier' and 'KStA' (Kölner Stadtanzeiger) are examples of large regional and local news outlets, while 'Berlin' is the information platform hosted by the Berlin municipal government. All of those sites provide high quality,
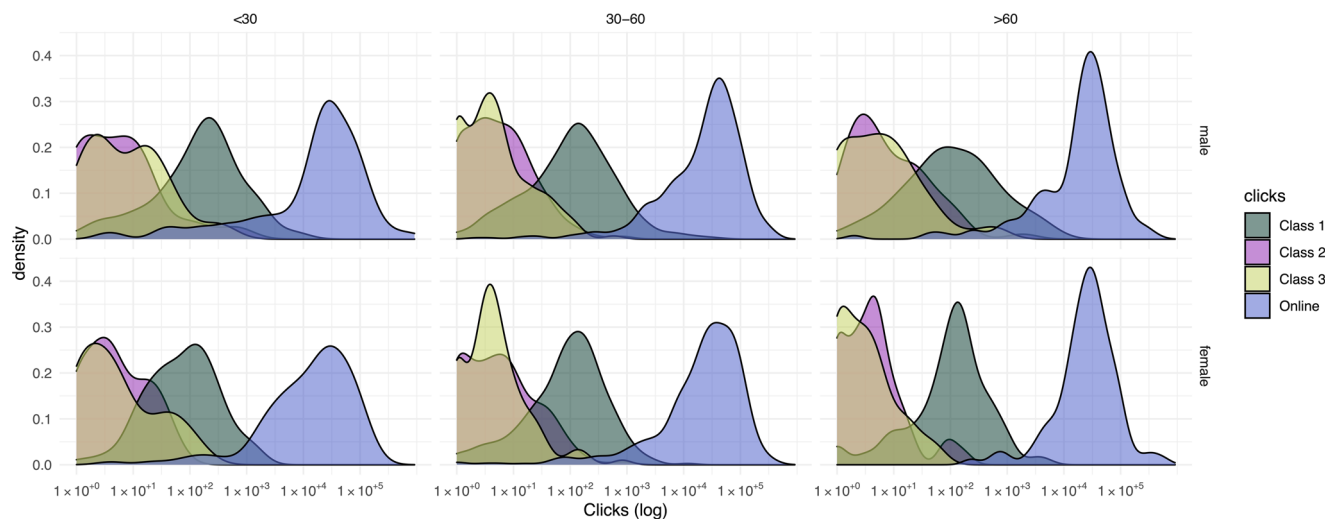
Fig. 4 Engagement with different classes of sites by age group and gender. Class 1: 'mainstream hubs', class 2: 'quality information providers' and class 3: 'niche forums'. 'Online' includes engagement with any websites recorded by the browser plug-in, including any political and a-political website visits.

journalistically-curated information, often with specific local focus, but apparently, neither do they offer extensive possibilities for political discussion, nor do they attract attention from diverse audiences.

Finally, the class of *niche forums* (class 3) contains websites with rather low conditional probabilities of fulfilling explicit criteria of deliberative potential, except for the potential of political expression and reciprocity in communication, and potential for political organization. In this class, rather niche online forums for specific communities, as well as forums that are dedicated to specific topics like gaming, cooking or anime content were found. While many domains in this class do not appear politically relevant at first glance, it is important to note that a manual validation step was taken to establish whether political discussion were indeed taking place on those websites. Examples of websites in class 3 are an esoteric forum that vividly discussed the upcoming federal elections, computer forums in which discussions on the military intervention in Afghanistan were found, a forum for children's second hand clothing ('Mamikreisel') and a forum dealing with issues of unemployment ('Eloforum') that hosted, partly in-depth, political discussions in niches of the forum.

In total, 34 websites were assigned to the mainstream hubs (class 1), 20 belong to the quality information providers (class 2) and 25 to the class of niche forums (class 3)[3]. The estimated mixing proportions corresponding to the share of observations belonging to each latent class are 49% for the mainstream hubs, 22% for the information providers and 29% for the niche forums.

The input criteria of deliberative potential form two natural groups: information, communication and participation are criteria that were coded manually and belong to the supply side of a website whereas connectivity, inclusivity and heterogeneity are coded computationally based on usage characteristics. This fundamental distinction is also reflected in the correlation-matrix between criteria. Therefore, the clustering process was repeated separately for the two groups of criteria (see SI C9 and C11). For the computationally-coded, demand-side criteria, a simple two factor solution was suggested with one class including all websites with high probabilities of fulfilling each criterion and one class with overall very low scores for connectivity, inclusivity and heterogeneity – in other words, high and low engagement websites. The model including only the manually-coded infrastructural criteria of information, communication and

participation possibilities suggested a more interesting pattern that is in line with the findings from the main model including all criteria. A first class contains websites with an strong information profile, including all public broadcasting pages. A second class contains websites with an especially strong forum component or communication profile with pages that also enable participation to some extent. The last class is rather a residual class including websites with overall low probabilities of fulfilling any criteria. The overall pattern largely mirrors the findings from the main model, the difference being that the two meaningful classes of the infrastructural model also contain the highly popular mainstream hubs that are, in the main model, separated through distinct patterns in the engagement based metrics. The latent class structure of the main model using all criteria was robust to the inclusion of alternative input variables, such as users' household income as feature of inclusivity and the size of the website, measured by the number of clicks as separate variable (see SI Fig. C12 and C13).

Engagement with different classes of sites. The measurement framework for the assessment of the deliberative potential of websites could, of course, be applied to various contexts for analytical and practical purposes. As one application, simple user-level engagement patterns, measured in the number of website visits as well as the duration of engagement is considered.

Given the underlying latent structure of deliberative potential dimensions, it does not surprise that the mainstream hubs are more frequently[4] accessed than quality information providers and niche forums (see Fig. C1a). However, if engagement is measured as duration instead of clicks, the engagement distributions become more similar (see Fig. C1b). This implies that people often access prominent websites like Google and Facebook in political contexts but that they tend to spend more time on public broadcasting platforms as well as small online forums to read news more carefully and, potentially, discuss political issues in depth within more tightly-knit communities compared to major social media platforms. More specifically, the duration per click ratio is only 35 s for mainstream hubs, on average, but 48 s for niche forums and almost a minute (59 s) for quality information providers. If the data were to be split, for example, just into news websites and social media platforms, this pattern would not have been observed (see Fig. C2a and C2b).

Another application is to switch from the perspective of the 'supply side' characteristics to the 'demand side' characteristics, namely the demographics of users engaging with different classes of sites. Figure 4 and C3 summarize the engagement with different classes of sites for different genders, age groups and levels of formal education. Despite some minor, though intuitive tendencies (e.g., the engagement with quality information providers is stronger than the engagement with niche online forums in the subgroup with the highest level of formal education (Abitur) in Germany) there is no clear pattern of selection visible within subgroups according to those three rough demographic indicators. The exploration of more sophisticated variables such as political orientation, political efficacy or political knowledge as possible driving factors for the selection into engaging with political content online remains subject to a subsequent project.

Discussion

The deliberative nature of an online environment is, as Beauchamp (2020) puts it, a function of membership and structures. In order to examine this function empirically, as a first step, this project systematically mapped the deliberative potential of those structures for the online public sphere in Germany. While this study is descriptive in nature, it is important to understand how increasingly complex media environments are composed of different arenas with different potential functions for democracy. While most of the previous research focused on specific aspects, such as online news media diets or the content of discussions on social media platforms (e.g., Esau et al. 2021; Guess 2021), this study took a step back and examined the infrastructure and usage patterns as the basis for online deliberation.

This study is one attempt - of probably many imaginable strategies - to map characteristics of a deliberative system empirically, that aimed to build closely onto the literature, by selecting and operationalising six deliberative criteria, in one political context, the German political online ecosystem. The resulting latent class structure is the result of this analytical strategy taken but not the 'ground truth' structure of a deliberative system that should from now on be applied to other media systems or even to the German political online ecosystem captured at another point in time. It is an empirical snapshot with the purpose to complement theoretical advancements with empirical observations. While the criteria structure is theoretically informed and could be applied to other contexts, the latent class structure, together with its engagement structure will look different across time and political context, for example, more partisan media systems like the United States.

While political online engagement only makes up a small proportion (about 1% of website visits) of the overall online engagement in Germany, a large part of the sample (1190 out of 1282) did engage with some political topics at least at some point around the federal elections in 2017. It is worth noting that the website selection approach, including a strict manual cross validation of whether a website actually featured political content, focuses on the minimization of false positives rather than false negatives. This implies quite a strict definition of 'politically relevant' and tends to rather underestimate the prevalence of political engagement online. However, possibly to the disappointment of many social scientists, engagement with political content online is by no means the dominant form of engagement.

The results of the study clearly align with Guess (2021) who found a considerable overlap of news media diets within a US sample that goes against the common notion of selective exposure in online 'echo chambers'. According to Guess (2021), this overlap originates from individuals' common use of large

mainstream hubs for political information. Correspondingly, in this German sample, the largest cluster of websites are highly-frequented sites that are commonly visited by a large proportion of users. These informational hubs can be understood to be a kind of general-interest intermediary that may indeed facilitate a common arena within the digital public sphere that offers shared experiences and the possibility of incidental encounters with diverse perspectives (Sunstein 2018).

The results of the latent class analysis further suggest that public service broadcasting still plays a major role in the German online public discourse even though these websites did not reach a particularly diverse audience within the sample. This finding aligns with previous work on deliberative democracy that, when mapping the television news ecosystem, identified an elite focused coverage within German public broadcasting (Wessler and Rinke 2014) which, however, speaks against the often implied view that public broadcasting is in itself lowering audience polarization through broad appeal. Furthermore, the reference to local or regional issues and information is a commonality of many websites assigned to the class of quality information providers. The local nature of political issues is often neglected when studying political online communication or when using digital trace data that do not have a geospatial component. However, on an interesting side note, Ellger et al. (2021) find that the decline of local newspapers can be related to an increase in political polarization, a relationship that could be given more attention in the study of online politics. While digital technology lets information flows transcend physical constraints, people still live in specific local contexts.

Furthermore, the analysis highlights a latent class of websites that is only mentioned in a small proportion of empirical studies on online deliberation. Wright (2012) coined the term 'third spaces' for non-political online spaces where political talk emerges based on case studies, similar to Graham (2012). This study demonstrates the importance of their early observations on a much larger basis. The class contains mostly niche forums dedicated to specific topics and communities which points to the phenomenon of incidental exposure to political issues online (Valeriani and Vaccari 2016; Yadamsuren and Erdelez 2010). Furthermore, these online communities might be comparably more tightly knit because of shared (apolitical) interests and fewer overall user numbers, which allows individuals to recognize each other (despite usually being pseudonymous, Moore et al. 2020). These forums, which, in comparison to large social media platforms, might be closer to offline social groups in which a basic form of trust can be established between members, can provide interesting possibilities for informal political discussions among citizens and might operate as important 'weak ties' between large online information and communication platforms within a deliberative online system (Esau et al. 2017; Granovetter 1973; S. W. Rosenberg 2014).

As visible among the mainstream hubs, website popularity is heavily ensconced in the three additional dimensions of deliberative potential (connectivity, inclusivity and heterogeneity). One obvious reason for this finding is that the degree of centrality of a node in a social network increases with the frequency of its interactions. Another measurement related explanation could be that the Shannon-Wiener diversity index puts more weight on richness than on evenness (Zeleny 2021), implying a rising index with more users. Therefore, caution must be taken against a substantive interpretation of the finding that the most heavily used platforms in the sample are, according to the measures, also the most 'inclusive' and 'heterogeneous'. While they are indeed a common source of information and a common arena of political communication for citizens with different demographic profiles and heterogeneous political attitudes, it is still important to keep

in mind that this does not prevent the formation of niche corners and sub-groups that might not speak to each other.

Another limitation to consider is that when classifying the content of the sites as political or not, the full URL-string was considered. While this often features the most important key-words of the page accessed, scraping the entire HTML text of the site might have been helpful in some cases[5].

The manually selected set of keywords naturally comes with certain boundary conditions. It is systematically easier to rigorously identify specific political terms, such as the names of politicians and terms referring to party politics and administrative processes in comparison to political issues like education and social policy because terms like 'family' or 'housing' appear in many different political and apolitical contexts. Various efforts were taken to reduce this imbalance as much as possible (see SI D).

Finally, the web tracking data is based on desktop use and does not include mobile devices. This certainly overlooks parts of users' political online engagement and may even introduce non-random blind spots. Furthermore, due to the temporal asymmetry between the browser histories (collected in 2017) and the content analysis on the respective websites (conducted in 2021), one cannot rule out the possibility that some websites might have changed in terms of structure, content and function for online public discourse.

One may ask which websites show the highest deliberative potential but this study explicitly avoids a summative ranking as the core of a systemic understanding implies that different arenas can fulfill different functions for public discourse (Bächtiger and Parkinson 2019). This study suggests that the empirical reality maps this normative account. Given that deliberative theory is fundamentally normative, one may consider possible normative implications for online public discourse that follow from this empirical mapping of a deliberative system. Certain combinations of deliberative potential criteria, such as the provision of communicative spaces that are characterized as inclusive and heterogeneous or the provision of high-quality political information in spaces with high connectivity to other relevant sources, clearly appear as normatively desirable (Mansbridge et al. 2012). However, a distinction between websites that primarily provide information and other websites that specialize on discussions, seems hardly detrimental to public discourse. On the contrary, this distinction could reflect the ideal of a shared factual baseline that is built by quality information providers on which basis then conflicting discussions can safely occur in other arenas (Habermas 2021; Krause 2008).

Accordingly, this study shows that few websites fulfill all criteria and some combinations of deliberative criteria are more frequent than others: information providing infrastructure often comes with high usage, reflected in heterogeneity and inclusivity, while communication also occurs in niches. Furthermore, in previous accounts theoretically distinct classes of websites, such as major newspapers and social media platforms, empirically sort into the same class when focusing on affordances and usage. However, the outlined systemic understanding that one website does not have to serve all criteria and the empirical findings about skewed participation in public discourse may allow a hypothesis about the critical state of the online media system: perhaps one website *should* also not try to serve all criteria. For example, public broadcasting and established newspapers are the backbone of quality information providence in Germany. Their increasing presence on social media, on the one hand, perhaps reaches otherwise lost audiences but on the other hand, risks eroding their core function of quality information providence that serves as common factual baseline for deliberation (Habermas 2021) through constraints imposed by the structure of social media. Visible engagement in the comment sections showcases the opinions and rhetoric of a skewed minority while for the largely silent majority that becomes visible in this study, public broadcasting remains a core provider of quality political information. Moreover, entering the market of digital content creators and advertisers is a competition that public broadcasting in Germany would not even have to play, given public funding combined with independent agenda setting.

Conclusion

This project illustrates that the internet provides a plethora of sources for political information, arenas for political communication and some opportunities for online participation. This study clearly found *potential* for public deliberation in the German speaking web in 2017. Even though political content is only a small proportion of the overall content accessed online—the German deliberative system seems to be a rather small fraction of the wider online environment—almost everyone in the sample engaged with some political content around the federal election in 2017. This implies that the consumption of political content is not as exclusive as the visible discussion patterns of few very active users on social media may imply.

The infrastructure of a deliberative system goes far beyond news websites and social media platforms but includes a wide range of different types of popular and niche platforms with different primary functions. On some platforms, users get political information. However, it is not clear if those are accurate or misinformation. On other platforms, they can discuss political issues, deliberatively or not. While only very few websites in the sample offer possibilities for participation, the demand also seemed limited.

Mainstream hubs are most central in the network of topical links, whereas public broadcasting outlets and especially the niche forums are more at the periphery of the network. Considering the definition of links within the connectivity measure, this implies that users move beyond the quick bites of political information on mainstream platforms but read more on the topic elsewhere. Those platforms appear to act as general-interest intermediary that may indeed facilitate a common arena within the digital public sphere that, against the notion of online 'echo chambers', offers shared experiences and the possibility of encounters with diverse perspectives. This finding aligns with the current state of the literature, finding limited empirical support for the prevalence and impact of online 'echo chambers' (e.g. Flaxman et al. 2016; Guess et al. 2023; Guess et al., 2021; Dubois and Blank 2018). The class of information providers can be interpreted as evidence for the persisting centrality of high quality public broadcasting as the backbone for democratic deliberation in Germany. The question as to whether we stand at the beginning or the end of the public broadcasting era online could be determined using detailed information on the user base. This project demonstrated the presence of a-political spaces in which political discussion emerges on a large empirical basis. While niche online forums are especially characteristic for the earlier years of the internet, it will be interesting to see in which spaces more tightly knit online communities will form in the future as previous exchange around a-political shared interests may build mutual trust as important basis for the discussion of conflicting political views.

Even though the found latent class structure appears intuitive, this structure was war from obvious as previous theoretical accounts have rarely moved beyond an assumed a split between news media and social media, a cyclographic split that was fed forward into empirical studies. Furthermore, the results of this study do reveal several surprising aspects. First, negative findings on the deliberative quality online are contrasted by findings about the potential of the political online ecosystem when examining passive audiences in contrast to digital traces of active social

media commenters. Second, the absence of central websites with low heterogeneity aligns with Guess et al. (2021) but provides more evidence against the otherwise common notion of online "echo chambers" (Sunstein 2002). Third, public broadcasting stood out as distinct class in a data driven, bottom-up approach, even with a sole focus on infrastructural elements and usage characteristics.

While this project empirically mapped the online media structures underlying online deliberation for the first time, the logical next step in the research agenda is the quantitative description of membership, the profiles of internet users engaging with political information and communication online. In particular because online political deliberation itself may not be a mainstream behavior, the mechanisms of selection into the online public discourse need to be determined.

Data availability

Notes

1 An alternative model including participants' household income as additional input criterion for inclusivity is reported in SI C13. In this German survey, ethnicity as another statistical marker of minority status was not asked.

2 The distribution of the self-reported political orientation of the sample approaches a normal distribution and also geographically, online activity patterns in the sample distribute about evenly across Germany.

3 The order of classes has no deeper meaning but is determined by configurations in the estimation process.

4 Cumulative engagement measures are baseline corrected, meaning that they show the share of website visits that users spend on e.g. quality information providers in relation to their total number of website visits in the measurement time frame.

5 However, an extremely robust scraper would have to be built in order to process hundreds of thousands of different domain structures (in the original full dataset). Future projects may try to build such a scraper, web-scrape all the sites and search for political topics in the full HTML text of websites instead of the URL-strings. The reference body (sites explicitly dealing with the 2017 German public discourse that I selected to generate keywords, see SI D) would then be similar enough to the target body (now being the full-text of websites instead of URL-text only) to use the semi-automated keyword extraction method proposed by King et al. (2017).

References

Anderson AA, Brossard D, Scheufele DA, Xenos MA, Ladwig P (2014) The "nasty effect:" Online incivility and risk perceptions of emerging technologies [Publisher: Oxford University Press Oxford, UK]. J computer-mediated Commun 19(3):373–387

Auguie B, Antonov A (2017) gridExtra: Miscellaneous Functions for "Grid" Graphics. Retrieved December 22, 2021, from https://CRAN.R-project.org/package= gridExtra

Bacher, J, Vermunt, JK (2010) Analyse latenter Klassen. In C Wolf & H Best (Eds.), Handbuch der sozialwissenschaftlichen Datenanalyse. VS Verlag für Sozialwissenschaften (pp. 553–574). Retrieved September 27, 2021, from https://doi.org/10.1007/978-3-531-92038-2_22

Bächtiger A, Parkinson J (2019) Mapping and measuring deliberation: Towards a new deliberative quality. Oxford University Press

Bail CA (2021) Breaking the Social Media Prism. Princeton University Press

Bakshy E, Messing S, Adamic LA (2015) Exposure to ideologically diverse news and opinion on Facebook. Science 348(6239):1130–1132

Barberá P, Jost JT, Nagler J, Tucker JA, Bonneau R (2015) Tweeting from left to right: Is online political communication more than an echo chamber? Psychol Sci. 26(10):1531–1542

Basu S (1999) Dialogic ethics and the virtue of humor. J Political Philos 7(4):378–403. https://doi.org/10.1111/1467-9760.00082

Beauchamp N (2020) Modeling and Measuring Deliberation Online. In The Oxford Hand-book of Networked Communication

Best for Planning (2017) Berichtsband b4p 2017. Gesellschaft für integrierte Kommunikationsforschung

Boswell J, Corbett J (2017) Why and how to compare deliberative systems. Eur J Political Res 56(4):801–819

Boxell L, Gentzkow M, Shapiro JM (2017) Greater Internet use is not associated with faster growth in political polarization among US demographic groups. Proc Natl Acad Sci 114(40):10612–10617

Bright J (2018) Explaining the emergence of political fragmentation on social media: The role of ideology and extremism. J Computer-Mediated Commun 23(1):17–33

Bright J, Bermudez S, Pilet J.-B, Soubiran T (2020) Power users in online democracy: Their origins and impact. Inf Commun Soc 23(13):1838–1853. https://doi.org/10.1080/1369118X.2019.1621920

Bright J, Marchal N, Ganesh B, Rudinac S (2022) How Do Individuals in a Radical Echo Chamber React to Opposing Views? Evidence from a Content Analysis of Stormfront. Hum Commun Res 48(1):116–145. https://doi.org/10.1093/hcr/hqab020

Buhl F, Günther E, Quandt T (2018) Observing the dynamics of the online news ecosystem: News diffusion processes among german news sites. J Stud 19(1):79–104

Carpini M. X. D, Keeter S (1993) Measuring Political Knowledge: Putting First Things First. Am J Political Sci 37(4):1179–1206. https://doi.org/10.2307/2111549

Cinelli M, de Francisci Morales G, Galeazzi A, Quattrociocchi W, Starnini M (2021) The echo chamber effect on social media. Proc Natl Acad Sci USA 118(9). https://doi.org/10.1073/pnas.2023301118

Coe K, Kenski K, Rains SA (2014) Online and uncivil? Patterns and determinants of incivility in newspaper website comments. J Commun 64(4):658–679

Coleman S, Moss G (2012) Under construction: The field of online deliberation research

Conover PJ, Searing DD, Crewe IM (2002) The Deliberative Potential of Political Discussion. Br J Political Sci 32(1):21–62. https://doi.org/10.1017/S0007123402000029

Cota W, Ferreira S, Pastor-Satorras R, Starnini M (2019) Quantifying echo chamber effects in information spreading over political communication networks. EPJ Data Sci 8(1). https://doi.org/10.1140/epjds/s13688-019-0213-9

Csardi G, Nepusz T (2006) The Igraph Software Package for Complex Network Research. Inter J Complex Syst 1695:1–9. http://igraph.org

Dahl DB (2016) Xtable: Export Tables to LaTeX or HTML

Díaz E, Koutra C (2013) Evaluation of the persuasive features of hotel chains websites: A latent class segmentation analysis. Int J Hosp Manag 34:338–347

Dixon P (2003) VEGAN, a package of R functions for community ecology. J Veg Sci 14(6):927–930

Döring M (2021) How-to Bureaucracy: A Concept of Citizens' Administrative Literacy. Administration & Society, 0095399721995460. https://doi.org/10.1177/0095399721995460

Douglas KM, Sutton RM, Cichocka A (2017) The Psychology of Conspiracy Theories. Curr Directions Psychol Sci. 26(6):538–542. https://doi.org/10.1177/0963721417718261

Dryzek JS (2012) Foundations and frontiers of deliberative governance. Oxford University Press

Dubois E, Blank G (2018) The echo chamber is overstated: The moderating effect of political interest and diverse media. Inf Commun Soc 21(5):729–745

Eady G, Nagler J, Guess A, Zilinsky J, Tucker JA (2019) How many people live in political bubbles on social media? Evidence from linked survey and Twitter data. Sage Open 9(1):2158244019832705

Ellger F, Hilbig H, Riaz, S, Tillmann P (2021) Local Newspaper Decline and Political Polarization

Elstub S, Ercan SA, & Mendonça RF (2019) Deliberative systems in theory and practice. Routledge

Ercan SA, Hendriks CM, Boswell J (2017) Studying public deliberation after the systemic turn: The crucial role for interpretive research. Policy Politics 45(2):195–212

Esau K, Fleuß D, Nienhaus S-M (2021) Different arenas, different deliberative quality? Using a systemic framework to evaluate online deliberation on immigration policy in Germany. Policy & Internet 13(1):86–112

Esau K, Friess D, Eilders C (2017) Design Matters! An Empirical Analysis of Online Deliberation on Different News Platforms. Policy Int 9(3):321–342. https://doi.org/10.1002/poi3.154

Federal Agency for Civic Education (2017) Jahresrückblick 2017: Bpb. Bundeszentrale für politische Bildung. https://www.bpb.de/politik/hintergrund-aktuell/261923/rueckblick-2017

Feezell JT (2018) Agenda setting through social media: The importance of incidental news exposure and social filtering in the digital era. Political Res 71(2):482–494

Felicetti A, Niemeyer S, Curato N (2016) Improving deliberative participation: Connecting mini-publics to deliberative systems. Eur Polit Sci Rev 8(3):427–448

Fishkin JS (2018) *Democracy When the People Are Thinking P: Revitalizing Our Politics Through Public Deliberation*. Oxford University Press

Fishkin JS, Senges M, Donahoe E, Diamond L, Siu A (2018) Deliberative polling for multistakeholder internet governance: Considered judgments on access for the next billion. Inf Commun Soc 21(11):1541–1554

Flaxman S, Goel S, Rao JM (2016) Filter Bubbles, Echo Chambers, and Online News Consumption. Public Opin Q 80(S1):298–320. https://doi.org/10.1093/poq/nfw006

Fletcher R, Nielsen R (2018) Are people incidentally exposed to news on social media? A comparative analysis. New Media Soc 20(7), 2450–2468. https://doi.org/10.1177/1461444817724170

Fleuß D, Helbig K, Schaal GS (2018) Four parameters for measuring democratic deliberation: Theoretical and methodological challenges and how to respond. Politics Gov. 6(1):11–21

Follesdal A (2010) The place of self-interest and the role of power in the deliberative democracy. J Political Philos 18(1):64–100

Freudenthaler R, Wessler H (2022) Mapping Emerging and Legacy Outlets Online by Their Democratic Functions—Agonistic, Deliberative, or Corrosive? [Publisher: SAGE Publications Inc]. Int J Press/Politics 27(2):417–438. https://doi.org/10.1177/19401612211015077

Galston WA (2001) Political knowledge, political engagement, and civic education. Annu Rev Political Sci 4(1):217–234

Gerber M, Bächtiger A, Shikano S, Reber S, Rohr S (2018) Deliberative abilities and influence in a transnational deliberative poll (EuroPolis). Br J Political Sci. 48(4):1093–1118

Grafton RQ, Nelson HW, Lambie NR, Wyrwoll PR (2012) Shannon–Wiener diversity index. Edward Elgar Publishing Limited, https://www.elgaronline.com/view/nlm-book/9781849803878/c_3011.xml

Graham T (2008) Needles in a haystack: A new approach for identifying and assessing political talk in nonpolitical discussion forums. Javn.- public 15(2):17–36

Graham T (2012) Beyond "political" communicative spaces: Talking politics on the Wife Swap discussion forum. J Inf Technol Politics 9(1):31–45

Granovetter MS (1973) The strength of weak ties. Am J Sociol 78(6):1360–1380

Grönlund K, Herne K, Setälä M (2015) Does enclave deliberation polarize opinions? Political Behav 37(4):995–1020

Guerrero-Solé F, Lopez-Gonzalez H (2019) Government Formation and Political Discussions in Twitter: An Extended Model for Quantifying Political Distances in Multiparty Democracies. Soc Sci Comput Rev 37(1):3–21. https://doi.org/10.1177/0894439317744163

Guess AM (2021) (Almost) Everything in Moderation: New Evidence on Americans' Online Media Diets. Am J Political Sci 65(4):1007–1022. https://doi.org/10.1111/ajps.12589

Guess AM, 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 YM, Lazer D, Moehler D, Nyhan B, Rivera CV, Settle J, Thomas DR, Tucker JA (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

Habermas J (1984) *The Theory of Communictive Action*. Beacon Press

Habermas J (1996) *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy (1st edition)*. Polity

Habermas J (2021) Überlegungen und Hypothesen zu einem erneuten Strukturwandel der politischen öffentlichkeit. In *Ein neuer Strukturwandel der öffentlichkeit?* (pp. 470– 500). Nomos Verlagsgesellschaft mbH & Co. KG

Habermas J, Lennox S, Lennox F (1974) The public sphere: An encyclopedia article (1964). N Ger Crit 3:49–55

Halpern D, Gibbs J (2013) Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. Comput Hum Behav 29(3):1159–1168

Huling J (2019) Jcolors: Colors Palettes for R and 'ggplot2', Additional Themes for 'ggplot2'. Retrieved December 22, 2021, from https://CRAN.R-project.org/package=jcolors

Janssen D, Kies R (2005) Online forums and deliberative democracy. Acta política 40(3):317–335

Jefferson T (1789) Thomas Jefferson to Richard Price - Thomas Jefferson | Exhibitions Library of Congress. Retrieved July 6, 2021, from https://www.loc.gov/exhibits/jefferson/60.html

Jensen JL (2003) Public spheres on the Internet: Anarchic or government-sponsored–A comparison. Scand Political Stud. 26(4):349–374

Jonsson ME (2015) Democratic innovations in deliberative systems–the case of the Estonian citizens' assembly process. J Public Délib. 11(1):7

Keyes O, Jacobs J, Schmidt D, Greenaway M, Rudis B, Pinto A, Khezrzadeh M, Meilstrup P, Costello A, Bezanson J et al. (2019) *Urltools: Vectorised tools for URL handling and parsing*

Kiernan D (2014) Quantitative Measures of Diversity, Site Similarity, and Habitat Suitability. In *Natural Resources Biometrics*. Open SUNY Textbooks

Kim JW, Guess A, Nyhan B, Reifler J (2021) The distorting prism of social media: How self-selection and exposure to incivility fuel online comment toxicity. J Commun 71(6):922–946

Kim Y, Chen H-T, De Zúñiga HG (2013) Stumbling upon news on the Internet: Effects of incidental news exposure and relative entertainment use on political engagement. Computers Hum Behav 29(6):2607–2614

King G, Lam P, Roberts M (2017) Computer-Assisted Keyword and Document Set Discovery from Unstructured Text. Am J Political Sci 61(4):971–988

Koiranen I, Koivula A, Keipi T, Saarinen A (2019) Shared contexts, shared background, shared values – Homophily in Finnish parliament members' social networks on Twitter. Telemat Inform 36:117–131. https://doi.org/10.1016/j.tele.2018.11.009

Krause SR (2008) *Civil passions*. Princeton University Press

Kriplean T, Toomim M, Morgan J, Borning A, Ko AJ (2012) Is this what you meant? Promoting listening on the web with reflect. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1559–1568. https://doi.org/10.1145/2207676.2208621

Lelkes Y (2016) Mass polarization: Manifestations and measurements. Public Opin Q 80(S1):392–410

Lelkes Y (2020) A bigger pie: The effects of high-speed Internet on political behavior. J Computer-Mediated Commun 25(3):199–216

Lin TH, Dayton CM (1997) Model selection information criteria for non-nested latent class models. J Educ Behav Stat 22(3):249–264

Linzer DA, Lewis JB (2011) poLCA: An R package for polytomous variable latent class analysis. J. Stat. Softw. 42(10):1–29

Long JA (2021) Panelr: Regression Models and Utilities for Repeated Measures and Panel Data. Retrieved December 22, 2021, from https://CRAN.R-project.org/package=panelr

Lorenz-Spreen P, Oswald L, Lewandowsky S, Hertwig R (2021) Digital Media and Democracy: A Systematic Review of Causal and Correlational Evidence Worldwide

Maia RCM (2017) Politicization, New Media, and Everyday Deliberation. In P Fawcett, M Flinders, C Hay, & M Wood (Eds.), *Anti-Politics, Depoliticization, and Governance (p. 0)*. Oxford University Press. https://doi.org/10.1093/oso/9780198748977.003.0004

Manosevitch E, Walker D (2009) Reader comments to online opinion journalism: A space of public deliberation. Int Symp Online Journalism 10:1–30

Mansbridge J, Bohman J, Chambers S, Christiano T, Fung A, Parkinson J, Thompson DF, Warren ME (2012) A systemic approach to deliberative democracy. In J Parkinson & J Mansbridge (Eds.), *Deliberative Systems*. Cambridge University Press (pp. 1–26)

Menchen–Trevino E (2013) Collecting vertical trace data: Big possibilities and big challenges for multi-method research. Policy Int 5(3):328–339

Mendonça RF, Ercan SA (2015) Deliberation and protest: Strange bedfellows? Revealing the deliberative potential of 2013 protests in Turkey and Brazil. Policy Stud 36(3):267–282

Mendonça RF, Ercan SA, Asenbaum H (2022) More than Words: A Multi-dimensional Approach to Deliberative Democracy. Political Stud 70(1):153–172. https://doi.org/10.1177/0032321720950561

Messner M, Distaso MW (2008) The source cycle: How traditional media and weblogs use each other as sources. Journalism Stud 9(3):447–463

Moore A, Fredheim R, Wyss D, Beste S (2021) Deliberation and Identity Rules: The Effect of Anonymity, Pseudonyms and Real-Name Requirements on the Cognitive Complexity of Online News Comments. Political Stud 69(1):45–65. https://doi.org/10.1177/0032321719891385

Munzert S, Ramirez-Ruiz S, Barberá P, Guess AM, Yang J (2021) Cheating in Online Assessments of Political Knowledge: Evidence from Survey and Digital Trace Data

Munzert S, Selb P (2017) Measuring political knowledge in web-based surveys: An experimental validation of visual versus verbal instruments. Soc Sci Comput Rev 35(2):167–183

Neuberger C, Nuernbergk C, Langenohl S (2019) Journalism as multichannel communication: A newsroom survey on the multiple uses of social media. Journalism Stud 20(9):1260–1280

Niemeyer S, Curato N, Bächtiger A (2015) Assessing the deliberative capacity of democratic polities and the factors that contribute to it. Democracy: A Citizens'Perspective, Abo (Turku), Finland

Nonnecke B, Preece J (1999) Shedding light on lurkers in online communities. *Ethno-graphic studies in real and virtual environments: Inhabited information spaces and connected communities, Edinburgh*, 123128

Ohlsen N (2015) Example for a latent class analysis with the poLCA-package in R – ahoi data. Retrieved June 28, 2021, from https://statistics.ohlsen-web.de/latent-classanalysis-polca/

Oksanen J (2013) *Vegan: Ecological diversity*. R Project, 368

Ooms J (2018) *Writexl: Export Data Frames to Excel 'xlsx' Format*. https://CRAN.Rproject.org/package=writexl

Pariser E (2011) *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin

Parkinson J, Mansbridge J (2012) *Deliberative systems: Deliberative democracy at the large scale*. Cambridge University Press

Pebesma E, Bivand R, Racine E, Sumner M, Cook I, Keitt T, et al. (2018). *Sf: Simple Features for R. R package version 0.6-1*

Pedrini S (2014) Deliberative capacity in the political and civic sphere. Swiss Political Sci Rev 20(2):263–286

Prior M (2005) News vs. entertainment: How increasing media choice widens gaps in political knowledge and turnout. Am J Political Sci 49(3):577–592

Quandt T (2018) Dark Participation. Media Commun 6(4):36–48. https://doi.org/10.17645/mac.v6i4.1519

R Core Team (2018) R: A language and environment for statistical computing

Rathje S, Van Bavel JJ, van der Linden S (2021) Out-group animosity drives engagement on social media. Proc Natl Acad Sci 118:26

Richardson JE, Stanyer J (2011) Reader opinion in the digital age: Tabloid and broadsheet newspaper websites and the exercise of political voice. Journalism 12(8):983–1003

Rivero G (2019) Preaching to the choir: Ideology and following behaviour in social media. Contemp Soc Sci 14(1), 54–70. https://doi.org/10.1080/21582041.2017.1325924

Rivers D (2006) Sample matching: Representative sampling from internet panels. *Polimetrix White Paper Series*

Rosenberg JM, Beymer PN, Anderson DJ, Van Lissa C, Schmidt JA (2019) tidyLPA: An R package to easily carry out latent profile analysis (LPA) using open-source or commercial software. J Open Source Softw 3(30):978

Rosenberg SW (2014) Citizen competence and the psychology of deliberation. In Elstub S, McLaverty P (eds), Deliberative democracy: Issues and cases (pp. 98–117). Edinburgh University Press

Rowe I (2015) Deliberation 2.0: Comparing the deliberative quality of online news user comments across platforms. J. Broadcasting Electron. media 59(4):539–555

Salem H, & Stephany F (2020) Wikipedia: A Challenger's Best Friend? Utilising Information-seeking Behaviour Patterns to Predict US Congressional Elections. *Util-ising Information-seeking Behaviour Patterns to Predict US Congressional Elections (October 27, 2020)*

Scudder MF (2020) *Beyond Empathy and Inclusion: The Challenge of Listening in Democratic Deliberation*. Oxford University Press

Steenbergen MR, Bächtiger A, Spörndli M, Steiner J (2003) Measuring political deliberation: A discourse quality index. Comp Eur Politics 1(1):21–48

Steyaert J (2000) Local governments online and the role of the resident: Government shop versus electronic community. Soc Sci Computer Rev 18(1):3–16

Stier S, Breuer J, Siegers P, Thorson K (2020) Integrating Survey Data and Digital Trace Data: Key Issues in Developing an Emerging Field. Soc Sci Computer Rev. 38(5):503–516

Strandberg K, Grönlund K (2018) Online Deliberation. In *The Oxford Handbook of Deliberative Democracy*. Retrieved October 22, 2021, from https://doi.org/10.1093/oxfordhb/9780198747369.001.0001/oxfordhb9780198747369-e-28

Strauss N, Huber B, Gil de Zuniga H (2020) "Yes, I Saw It - But Didn't Read It …" A Cross-Country Study, Exploring Relationships between Incidental News Exposure and News Use across Platforms. Digit Journalism 8(9):1181–1205. https://doi.org/10.1080/21670811.2020.1832130

Sun N, Rau PP-L, Ma L (2014) Understanding lurkers in online communities: A literature review. Computers Hum. Behav. 38:110–117

Sunstein CR (2002) The law of group polarization. J Political Philos 10(2):175–195

Sunstein CR (2018) *# Republic: Divided democracy in the age of social media*. Princeton University Press

Tewksbury D, Weaver AJ, Maddex BD (2001) Accidentally informed: Incidental news exposure on the World Wide Web. Journalism mass Commun. Q. 78(3):533–554

Thompson DF (2008) Deliberative democratic theory and empirical political science. Annu Rev Polit Sci 11:497–520

Ugarriza E, Caluwaerts D (2014) *Democratic Deliberation in Deeply Divided Societies: From Conflict to Common Ground*. Springer

Valeriani A, Vaccari C (2016) Accidental exposure to politics on social media as online participation equalizer in Germany, Italy, and the United Kingdom. N. Media Soc. 18(9):1857–1874

Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. Science 359(6380):1146–1151. https://doi.org/10.1126/science.aap9559

Wallsten K (2007) Agenda setting and the blogosphere: An analysis of the relationship between mainstream media and political blogs. Rev Policy Res. 24(6):567–587

Warnes GR, Bolker B, Gorjanc G, Grothendieck G, Korosec A, Lumley T, Mac-Queen D, Magnusson A, Rogers, J et al. (2015) *Gdata: Various R Programming Tools for Data Manipulation*. http://CRAN.R-project.org/package=gdata

Warren ME (2021) A bridge across the democracy–expertise divide

Wessler H, Rinke EM (2014) Deliberative Performance of Television News in Three Types of Democracy: Insights from the United States, Germany, and Russia. J. Commun. 64(5):827–851. https://doi.org/10.1111/jcom.12115

Wickham H (2016) Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J et al. (2019) Welcome to the Tidyverse. J. Open Source Softw. 4(43):1686

Wickham H, Bryan J (2018) *Readxl: Read Excel Files*. https://CRAN.R-project.org/package=readxl

Wickham H, Francois R (2015) Dplyr: A Grammar of Data Manipulation

Wickham H, Hester J, Francois R (2017) *Readr: Read Rectangular Text Data*

Wiklund H (2005) A Habermasian analysis of the deliberative democratic potential of ICT-enabled services in Swedish municipalities. N Media Soc 7(2):247–270

Wright S (2012) From "Third Place" to "Third Space": Everyday Political Talk in Non-Political Online Spaces [Publisher: Routledge _eprint. Javnost - The Public 19(3):5–20. https://doi.org/10.1080/13183222.2012.11009088

Xie Y (2018) Knitr: A comprehensive tool for reproducible research in R. Implementing reproducible research. Hall/CRC, Chapman, p 3–31

Yadamsuren B, Erdelez S (2010) Incidental exposure to online news. Proc Am Soc Inf Sci Technol 47(1):1–8

Yang T, Majó-Vázquez S, Nielsen RK, González-Bailón S (2020) Exposure to news grows less fragmented with an increase in mobile access. Proc. Natl Acad. Sci. 117(46):28678–28683

Young IM (2002) Representation and Social Perspective. In IM Young (Ed.), Inclusion and Democracy. Oxford University Press. https://doi.org/10.1093/0198297556.003.0005

Zeleny D (2021) Analysis of community ecology data in R. Retrieved October 11, 2021, from https://www.davidzeleny.net/anadat-r/doku.php/en:start

Zhu H (2019) KableExtra: Construct complex table with'kable'and pipe syntax. R. package version 1:0

Ziegele M, Quiring O, Esau K, Friess D (2020) Linking news value theory with online deliberation: How news factors and illustration factors in news articles affect the deliberative quality of user discussions in SNS'comment sections. Commun Res 47(6):860–890

Acknowledgements

Author contributions

LO contributed to the conceptions of the study, data preprocessing, data analysis as well as the preparation and revision of the manuscript.

Funding

Competing interests

The author declares no competing interests.

Ethical approval

The collection of survey and tracking data used in this study was approved by the IRBs of Princeton University (protocols 8327, 10014, and 10041) and the University of Southern California (UP-17-00513) and authorized by the University of Illinois via a designated IRB agreement.

Informed consent

Explicit and informed consent was obtained from all participants whose data was collected.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-024-03115-0.

Correspondence and requests for materials should be addressed to Lisa Oswald.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.