4. Social Media Platforms as Parts of Deliberative Systems

Social media platforms are highly relevant for individuals and societies under the conditions of the digital constellation.[1] They are a dominant infrastructure for communication under the conditions of the digital constellation. Digital platforms, more generally, connect people with each other, and they thus "*set the stage for actions* to unfold through ordered emergence" (Bratton 2016: 47, orig. emph.). It does not surprise, therefore, that the platforms and their logics also considerably shape the deliberative systems that operate under the conditions of the digital constellation.

The term platform has been a buzzword of academic literature in recent years – and has consequently been modified in various ways, including "platformisation" (Nieborg/Poell 2018; Schrape 2021b), "platform capitalism" (Srnicek 2017), "platform socialism" (Muldoon 2022) and "platform" combined with practically any other supplement. However, what is conveyed by this term varies widely. Diverse scholarly communities use it differently; for example, "there are subtle variations in the ways that platforms are understood by computer scientists, economists, digital media scholars, and lawyers" (Gorwa 2019), as well as by the companies providing platform services themselves (cf. Carrigan/Fatsis 2021: 33). Even among digitisation and social media researchers, its definition varies depending on who uses it and with what aim and focus it is used (cf. ibid.). But there are some central logics to social media platforms that most scholars agree on. I will first present these logics – which will form the basis for my analyses in chapters 5 and 6 – and I will then, in

1 The importance of social media platforms to society-wide political communication has already been alluded to during the previous chapter, and I will further elaborate on that in chapter 4.2.

the second part of this chapter, address the role and relevance of social media platforms for deliberative systems.

4.1 The Logics of Social Media Platforms

"A working technical definition of platform, in general, may include references to a standards-based technical-economic system that simultaneously distributes interfaces through their remote coordination and centralizes their integrated control through that same coordination" (Bratton 2016: 42). In the case of digital platforms, the foundational standards of such systems include the production, encoding, storing, distributing, decoding and consumption of information in binary code (cf. Jungherr et al. 2020: 7). The systems then mostly reify as

> hardware- and software-based, programmable and algorithmically structuring technological infrastructures […] through which information is exchanged, communication is organized, work and markets are coordinated, a broad spectrum of services is offered or digital and material products are distributed (Dolata/Schrape: 14).

They are thus "an infrastructure for processing that data (content, traces, patterns of social relations) for customer service and for profit" (Gillespie 2018c: 254). Acting as intermediaries, platforms facilitate transactions between third parties on the internet. They give access to, host, store, organise and transmit content, products and services provided by third parties on the internet (cf. OECD 2010; Gillespie 2018a: 19). Thus, they "link actors, information, and events across multiple spatial and temporal scales at once" (Bratton 2016: 50). While platforms do not produce contents themselves, they still shape the conditions in which users operate through their design and architecture (cf. Muldoon 2022: 26–27), determining what type of content is published, how it is seen and by whom. They are

> environments that have been painstakingly developed to guide users towards actions that are profitable for the company. Alongside explicit policies that regulate how parties will interact and on what terms, digital architects have built online landscapes that encourage certain types of behaviour through a deep understanding of human psychology and motivations built

> up over decades of social psychology and marketing studies. The seemingly smooth and open online space is in fact a carefully designed gradated and stratified world with gentle slopes that guide users towards specific sets of behaviours. (ibid.)

Consequently, the platform metaphor has been repeatedly criticised for obfuscating the active interventions and the value-adding performed by the services themselves.[2]

Many of the digital platforms that can be described in this way are social media platforms. The contents transmitted on social media platforms are limited to communication, rather than services. Platform users produce "content of their own"; they "see and interact with the content of their friends or other users online" (Balaban et al. 2019: 31). Users meet as essentially equal participants in a communicative exchange on topics of their own choice (cf. Habermas 2022: 159).[3]

The platforms structure and personalise this exchange for each individual user. In order to do this, social media platforms collect and *store* user data, and they *process* these data for a variety of different purposes, mainly to generate value for their shareholders (see chapter 4.2). Thus, platforms use a "research-centric apparatus", an "analytic architecture" (Marres 2018: 435) that is based on data regarding user behaviour.

4.1.1 Collecting, Standardising and Storing Data

The data may include not only data concerning the contents explicitly provided by the users, but usually also specifics about user behaviour on the platform itself and beyond. Platforms are "assembling data points that invariably result from even the most innocuous actions, and even inactions, that individuals take in everyday life" (Pistor 2020: 111–112). A user is not necessarily aware that her data are being harvested; she "may have been put on notice in the agreement that she clicks when accessing a platform, but the act of surveillance is not

2 Cf. Gillespie (2010, 2017), Muldoon (2022), Nieborg/Poell (2018: 4276). Nonetheless, I will stick with the term, since it is widely used – even by critical authors – to describe the procedures that I address in my analyses.

3 Thus, social media platforms cannot be classified as traditional media outlets that produce content for the mass market. Neither are they neutral transmission service providers, such as the postal service (cf. Schrape 2021a: 155, 2021b: 5–6; also: Habermas 2021: 498).

felt or heard" (ibid.). Thus, social media platforms collect and store a wide variety of data.[4] These data include the contents a user has posted[5] and the respective meta-data,[6] as well as "explicit feedback" data (cf. Cardon 2019) regarding the user's interaction with the contents of others,[7] other behavioural data[8] and information that has explicitly been given.[9] Moreover, platforms store "implicit feedback" data (cf. ibid.);[10] they gather information on a user's navigation on the Web and on their device even beyond the respective platform,[11] as well as meta-data on files that are stored on the device and granted access via the platform (cf. ibid.); and they store contextual data such as unique identifiers,[12]

4 The following enumerations in this paragraph and the respective details in the footnotes have been condensed from the detailed lists that can be found in Meta's Privacy Center (cf. Meta 2022a) for Facebook and Instagram, TikTok's privacy policy statement (cf. TikTok 2022) and Google's Privacy & Terms (cf. Google 2022) YouTube, as well as from Cardon (2019) and Rieder (2020).

5 These include comments, videos, audios, pictures, comments and direct messages. Platforms can store them independently of the user's actively uploading or saving them to the platform; they also store anything that has been loaded into the buffer memory in order to copy it to or anything from the platform. They analyse the contents on videos and pictures, identifying and storing data on objects and landscapes, faces or other body parts and text of what has been spoken on a video.

6 This term refers to data regarding when, where and by whom the contents have been produced.

7 This includes data such as clicks, reactions, sharing and comments and their respective meta-data (cf. Cardon 2017; Lada et al. 2021; Stark/Stegmann 2020: 6) – these interaction data provide information on all participating individuals and on those individuals that might be mentioned in the contents, as well as on the relationships between all these individuals (cf. Pistor 2020: 104–105).

8 This includes time span and frequency of platform use, user settings, query and search terms, consumed contents, interaction with ads, voice and audio information when assistants are used, purchase activities and data concerning the user's contacting of the platform when reporting violations against community standards and laws or when requesting technical support, as well as data from surveys, advertisement and marketing campaigns, competitions and events.

9 This includes name, date of birth email address, payment information, biography, profile picture or contact lists imported from the mobile phone.

10 This includes reading or scrolling speed, watch time, movements of the mouse and patterns and rhythms of keyboard typing.

11 This involves collecting referrer URLs and cookie data or plug-in data.

12 This can be for device, browser, app or user; it is done for tracking cross-device use of the platform.

the network through which the user enters the platform,[13] performance data[14] and data concerning the device,[15] as well as movement traces in navigational or physical spaces, device signals[16] and sensor readings (cf. Cardon 2019; Rieder 2020: 36).[17] They usually also have agreements with advertising, measurement and data partners to exchange these data, so that they can complete the user profiles.

Based on these data, the platforms generate user profiles – "*whether they are desired or not*" (Bratton 2016: 49, orig. emph.). These profiles reflect user behaviour on three levels: their knowledge-seeking (represented by data on their queries and data on navigation on the platform and beyond), their physical data (approximated trace data of physical navigation, sensors of the environment such as smart homes, sensor data on their bodies, e.g., from smart watches) and their social data (gathered from explicit feedback, friendships, group or list associations, subscriptions, etc.) (cf. Stalder 2016: 189–190). Of course, these profiles do not mirror the actual individuals, but only the information that is available to the platform (cf. Bail 2021: 7–8; Gillespie 2014: 173–174; Stalder 2016: 196).[18] Therefore, the platforms' structuring of communication processes cannot respond directly to the individuals involved but only to the datasets collected for their user profiles.

Not all the different types of data are collected and used by all social media platforms in the same way and with equal weight. Differences depend on

13 This includes mobile provider, language, time zone, mobile number, IP address, SIM card, transmission speed, data on other devices within the network and WiFi access points.

14 This includes crash reports, event logs and performance protocols.

15 This includes data on ID, model and age of the device, operating system, hardware and software, battery charge, signal strength, available memory, browser type and names and types of apps and data.

16 These include GPS, Bluetooth, WiFi, cell towers and beacons.

17 These include data collected via accelerometer or gyroscope.

18 These data – even those that constitute smaller amounts of publicly available data – can be and often are used by third parties as well. A Silicon Valley data analyst described an example of a company assessing credit scores: "with your social media profile, they know your name, they know the names of your friends, and they can tell if you're black or not. They can tell how wealthy you are, they can tell if you're a credit risk" (Tarnoff/Weigel 2020: 119). There is a "shadowy data industry that runs parallel to these major platforms – comprised of innumerable data brokers and other similar firms – [that] function as surveillance machines just as much as Mark Zuckerberg's platform does" (Srnicek 2020: 86).

the platform's aim, economic model, ethical standards and legal context. The content moderation systems of social media platforms mainly rely on data concerning the contents, connections, behaviour on the platform, devices and physical position of a user. In the past, platforms mostly based their moderation on explicit feedback data, but now they increasingly rely on implicit feedback data, too (cf. Cardon 2017: 140; Wei 2020). They gather extensive data on users, since otherwise, algorithmic moderation would not work as effectively (cf. Tarnoff/Weigel 2020: 112).

In order to process data, platforms need to standardise them (cf. Gillespie 2014: 170). The collected data must be "cleaned up" and "oriented" in a certain way before "facing" an algorithm (cf. ibid.: 170–171).[19] This is necessary, for instance, in order to make the data accessible for technical, formal and discursive protocols, interfaces and formats (cf. Bratton 2016: 44, 47), as well as for algorithmic systems. The processes of cleaning and orienting data include allocating data to specific categories or formats.[20] The specific modes of standardising and accessing data preform the possible ways that a platform can actually process these data. In fact, the different use cases of data processing are the main reason why platforms need data standardisation (cf. Stalder 2016: 182).[21]

19 Arguably, "'raw data' is an oxymoron" (Gitelman 2013), as data are not even "raw" while they are collected. "Both data and variables are always already 'cooked', meaning they have been generated by cultural operations and formed in cultural categories" (Stalder 2016: 193, author's translation).

20 The granularity and rigidity of standardisation in data storing has changed over time. "In the earliest database architectures, information was organized in strict and, as it turned out, inflexible hierarchies" (Gillespie 2014: 171). Since then, other modes of database management have been developed. Bernhard Rieder describes these new developments as following the notion of the fini-unlimited, whose central element is "the idea that a limited number of elements can yield an (almost) unlimited number of combinations or arrangements" (2020: 31).

21 Data standardisation mostly targets the properties level of digital technologies. But it can be argued that standardisations occur on the affordances and the social configurations level of the digital constellation as well. Technical protocols and platform community guidelines produce a "tendency on the part of some [social] systems and social processes to transform themselves according to the needs of the platforms that might serve and support them, both in advance of their participation with that platform and as a result of that participation" (Bratton 2016: 42). Moreover, this translates to the social configurations emerging under the conditions of the digital constellation. They "capture" activities by "channelling them through designed functionalities, interfaces, and data structures" (Rieder/Hofmann 2020: 2). For example, real-world political campaigns increasingly rely on digital social media platforms and accordingly adapt

4.1.2 Content Moderation and Recommender Systems

A central use case of data processing for social media platforms is content moderation, especially in recommender systems. "[R]anking, filtering, and recommending have become central techniques for facilitating the 'right' connections, whether between consumers and products, users and contents, or between people seeking interaction, friendship, or love" (Rieder/Hofmann 2020: 2). Content moderation is "the central commodity platforms sell". It is "meant to draw users in and keep them on the platform, in exchange for advertising and personal data" (Gillespie 2018b: 210). Algorithmic content moderation therefore is an integral part of the business model of social media platforms.[22]

There are several reasons why platforms moderate and why they do so with the help of algorithms. First, in order to actually generate profit, platforms sell advertisements that are presented to the user either as part of their personalised feeds or coupled with specific contents. Therefore, "[a]lgorithms for online content selection are designed to maximize […] what social media commentators stubbornly persist in calling 'engagement'" (Marres 2018: 430–431). Meanwhile, the sheer number of third parties that need to be connected make some kind of general, neutral way of connecting the third parties – for example, providing chronological or alphabetical indices of the entries and profiles – highly unpractical. In order to make sure that users continue using their services and in order to "present their best face to new users, to their advertisers and partners, and to the public at large" (Gillespie 2018a: 5) and to thus make

---

or reinvent their entire campaign strategies (cf. Maschewski/Nosthoff 2021), and certain standards must be kept when placing political ads on the platforms. Individuals get used to these behavioural standards as well as to the fact that digital technologies thus "take part in the social" (Dickel 2022: 21). So, due to the channelling of real-world situations and activities through the designed functionalities of digital platforms and the acceptance of these measures by society, the platform's inherent logics standardise real-world situations and social configurations.

22 Platforms also employ other modes of content moderation. For instance, they introduce frictions for posting less desirable contents like fake news, and they provide easier access to more desirable modes of expression such as one-click reactions. Moreover, they label certain types of posts, limit their reach or entirely remove them, and they can limit the visibility of certain groups or ban such groups from their services. All these types of moderation are at least partially performed by or based on the output of algorithms. I will come back to them throughout chapters 5 and 6.1.

money, platforms have to recommend contents an individual user might like and react to. While doing so, they also aim "to protect one user from another, or one group from its antagonists, and to remove the offensive, vile or illegal [content] to protect the users" (ibid.). Therefore, "platforms do, and must, moderate the content and activity of users, using some logistics of detection, review, and enforcement" (Gillespie 2018a: 21; for more detail, see Gorwa et al. 2020).

The most important moderation tool used by social media platforms is the algorithmic system. The "available data […] make algorithmic matching highly attractive" (Rieder/Hofmann 2020: 2). The datasets on user behaviour "have gotten large enough where you can start to consider variable interactions in a way that's becoming increasingly predictive" (Tarnoff/Weigel 2020: 112). Platforms therefore "process and analyze [datasets] with the goal of perfecting predictability about the future behavior of their targets" (Pistor 2020: 105).[23] Not all of these data are predictive in and of themselves, or directly after collection.

> [T]here are a number of problems where the actual individual variables themselves don't have a lot of meaning, or they are kind of ambiguous, or they are only very weak signals. There's information in the correlation structure of the variables that can be revealed, but only through really huge amounts of data. (Tarnoff/Weigel 2020: 112)

This information can be retrieved with the help of algorithms. More to the point, platforms *need* algorithms to retrieve the necessary information from the databases, since "[p]latforms' business models need a *scale* that is ungovernable by anything other than statistical techniques" (Ananny 2020: 364). The platforms "gain size and strength by mediating unplanned and perhaps even unplannable interactions" (Bratton 2016: 44), and these interactions are therefore so numerable that they can be managed only by algorithms.[24] In short, algorithmic content moderation is instrumental to the workings of social media platforms, and communication processes occurring on these platforms are fundamentally shaped by these moderation systems. Therefore, I will now present the central logics of the algorithmic moderation systems used by social media platforms, give an example of a recommender system

23 Especially large datasets may even enable predictions at the macro-level about aggregate behaviour, such as with GDP forecasts (cf. Pistor 2020: 105).

24 It is "inconceivable that all of the large-scale social media platforms will be broken up and become small enough for exclusively human curation and moderation" (Ananny 2020: 364).

and comment on the accessibility of these systems for scientific research. On this basis, throughout chapters 5 and 6, I will analyse the implications of these algorithmic systems for deliberative systems.

The term algorithm is often used as a "sloppy shorthand" (Bogost 2015) for complex data processing systems. From a computational perspective, however, the term captures the combination of logic and control. Logic is "the problem domain-specific component and specifies the abstract formulation and expression of a solution (what is to be done)", while control refers to "the problem-solving strategy and the instructions for processing the logic under different scenarios (how it should be done)" (Kitchin 2017: 16–17). So, in the first place, algorithms are merely a series of specific instructions on what to do, a "recipe" (Amoore 2020: 11).[25]

Within a digital platform, algorithmic systems can fulfil different tasks: "online search, social media filtering, linguistic translation, email spam filtering, content moderation, voice dictation, chaotic storage, route planning, fraud detection and content recommendation are *only* a few examples of what has become ubiquitous within platform firms" (Carrigan/Fatsis 2021: 46, orig. emph.). Arvind Narayanan (2023) distinguishes between algorithms that process contents and algorithms that propagate contents (see Table 2; p. 130).[26] These algorithms are, in fact, "a whole suite of algorithms, but they are tightly coupled" (Narayanan 2023); in other words, they are algorithmic systems.

25 Algorithms "need not be software: in the broadest sense, they are encoded procedures for transforming input data into a desired output, based on specified calculations" (Gillespie 2014: 167). Therefore, "the formulation of an algorithm is, in theory at least, independent of programming languages and the machines that execute them" (Kitchin 2017: 17). However, the algorithms and algorithmic systems used by social media platforms for predicting user behaviour and moderating communication are, of course, based on specific programming languages.

26 In this listing by Narayanan (2023), "content moderation" refers to the handling of contents that are unlawful or against the community guidelines. Though this conception of the term is fairly common, I will use content moderation in the more general sense of "ordering contents and matching them with users", since this is more compatible with the deliberative theory understanding of "moderation".

Table 2: Major Social Media Content Algorithms (Source: Narayanan 2023)

Major Social Media Content Algorithms	
Content processing	*Content propagation*
Face recognition Image filters Annotation (e.g. image tagging) Audio transcription Language translation Augmented & virtual reality …	Search Recommenation (feeds) Ad delivery and tergating Content moderation Friend recommendation Notification Trending …

For analysing deliberative systems, the most important algorithmic systems of social media platforms are recommender systems.[27] Recommender systems match contents with users at a given point in time. Based on these data pools, social media platforms perform algorithmic matching using two main logics of categorisation. The first logic uses datasets with predefined categories, and the second categorises the data dynamically with the help of machine-learning techniques. The general aim of both techniques is to present users with contents that they will engage with. In fact, the prediction of engagements represents the core of algorithmic systems (cf. Narayanan 2023).

To calculate the probabilities for engagement, algorithmic systems categorise users and posts by using similarity metrics, that is, by using data science techniques to calculate the similarity between users. As already mentioned, the platforms establish user profiles containing extensive sets of various kinds of standardised data and calculate similarities between users with regard to their networks, behaviour and demographics (cf. Narayanan 2023).[28] Simultaneously, algorithmic systems categorise the posts themselves,

27 However, they "may only be a small component in a much larger system that includes other various instances of ordering, ranging from data modelling to user-facing interfaces and functions that inform and define what users can see and do" (Rieder/Hofmann 2020: 5).

28 Network refers to followers/followees, subscriptions, commenting, etc.; behaviour similarity refers to whether or not users have engaged with a similar set of posts; demographics refers to attributes such as age, gender and, most importantly, language and geography (cf. Narayanan 2023). This last form of categorisation, in particular, is a highly political process that shapes a platform's possibilities for using these data: "For-

based on the posts' similarities regarding their content and their similarities regarding the users that have already engaged with the posts. In accordance with these categorisations, algorithmic systems predict a user's engagement with a post based on the behaviour of similar users with similar posts. In other words, "[a]s we rate and approve, upvote and downvote, we sort ourselves into clusters. Recommendation systems in turn feed us content similar to what is popular with 'people like me,' […]" (Silver 2021). This approach is called collaborative filtering (cf. Healy 2015: 191). With time, these categories have become ever more fine-grained and dynamic. The statistical methods used to analyse the data have been developed to become increasingly sophisticated, and the algorithmic systems they are part of have become increasingly complex.

In addition to these methods from statistics and data science, platforms have introduced machine-learning techniques into their systems. These techniques aim to predict user behaviour and to optimise their own prediction methods in the process (Tarnoff/Weigel 2020: 114). In consequence, "[y]ou can start to see new things emerge that would not emerge from more standard ways of looking at problems" (ibid.: 113). Machine-learning systems are highly dynamic; multiple data are factored in, and the platform companies and designers, content providers and users that react to these contents each contribute in specific ways; everything takes place in a specific cultural context and is shaped by certain social practices, and all these different factors interact with and influence each other (cf. Rieder/Hofmann 2020: 8).

In order to unravel this – and in order to prepare for more detailed analyses of the effects of these moderation systems on deliberative systems – it will be helpful to look at how Meta explained the basic logics of the Facebook News Feed algorithm. Though there may be specific differences between YouTube, TikTok or any other system that make personalised content recommendations to users, these broad logics apply to all of the major platforms (cf. Narayanan 2023).[29]

---

mats for gender, race, health, and credit specify the shapes of our data, be it via high-performance, machine-learning systems or legacy paper machines like punch-card indexes. These formats are political not only in the way that they function as political prostheses for classical political dynamics of coercion, but more significantly they are political in the way they serve to perform the work of […] 'fastening' subjects to all manner of databases and systems" (Koopman 2021: 4).

29 It is to be expected that these logics will be complemented by artificial intelligence models in the foreseeable future. Yet the main logics behind the matching on social media platforms will still apply.

Lada and other authors from Meta (2021) describe how the platforms seek to provide each user with individualised recommendations. The algorithmic system of Facebook's News Feed uses multiple machine-learning models to produce different predictions about what content on the platform a user might engage with. "Each model tries to rank these pieces of content for Juan [an exemplary user]. Sometimes they disagree […]. So we need a way to combine these varying predictions into one score" (ibid.). Broadly speaking, this integration happens in five steps.

> First, the system collects all the candidate posts we can possibly rank for Juan […]. This eligible inventory includes any post shared with Juan by a friend, Group, or Page he's connected to that was made since his last login and has not been deleted. (Lada et al. 2021)[30]

To this pool, the algorithmic system adds "posts that were ranked for Juan (but not seen by him) in his previous sessions" (unread bumping logic), as well as "posts Juan has already seen that have since triggered an interesting conversation among his friends" (action-bumping logic) (cf. Lada et al. 2021).[31] In a second step, "the system needs to score each post for a variety of factors, such as the type of post, similarity to other items, and how much the post matches what Juan tends to interact with" (ibid.). This is where the similarity metrics I mentioned above come into play. Each post is compared with posts the user has already interacted with, its similarity to these posts is assessed and the contents are clustered accordingly. In a third step, the pool is further narrowed down to approximately 500 posts and pre-ordered by some preliminary processes. The fourth step is the main scoring pass. Most of the personalisation happens in this process.

> Here, a score for each story [that is: post] is calculated independently, and then all 500 posts are put in order by score. For some, the score may be higher for likes than for commenting, as some people like to express themselves more through liking than commenting. Any action a person rarely engages

30 According to this description, the algorithmic systems only suggest contents from profiles that are already connected to the user. At a later point, Meta also introduced "recommended posts" from unconnected profiles into the feeds of Facebook and Instagram (cf. Newton 2022b).

31 This has implications for the temporal structures of communication processes (see chapter 6.1).

> in (for instance, a like prediction that's very close to zero) automatically gets a minimal role in ranking, as the predicted value is very low. (ibid.)[32]

So, the ordering here is largely weighted by predictions about the user's preferences regarding engagement with contents; these predictions are inferred from the user's past behaviour. With any action the user takes on the platform, the algorithm adapts the weights attributed to the respective actions. Thus, the system always "learns" more about the user and how to improve its recommendation, and it self-optimises accordingly. In other words, "[t]hese techniques develop decision models inductively […], based on an arrangement between data, feedback, and a given purpose" (Rieder/Hofmann 2020: 7). In a fifth step, the contextual pass is applied. In this pass, "contextual features like content type diversity rules are added to make sure Juan's News Feed has a good mix of content types and he's not seeing multiple video posts, one after another" (Lada et al. 2021). This diversity "is also a defense against the algorithm's uncertainty about what the user wants at any given moment, because even the best algorithm is far from perfect at predicting engagement" (Narayanan 2023).

In sum, the process followed by the algorithmic system behind Facebook's News Feed continually narrows down the pool of contents that might cause the user to engage, with the probability of an engagement being approximated by similarity metrics and statistics based on the user's past behaviour.[33] Such systems are complex and multi-layered, and my descriptions are merely highly

32 In chapters 5 and 6, the relevance and impacts of one-click reactions and quantified communication will be repeatedly addressed.

33 The recommendations that are produced by such systems are neither neutral nor reproducible; they are not even necessarily the best recommendations the user could have received. Or, put more poetically: "From the gigantic haystack of searchable information, results are generated that are immediately declared to be the needle. How these results come about […] is only rudimentarily explainable, at best. But as long as the needle is even moderately functional, most users are content, and the algorithm registers this contentment in order to validate itself" (Stalder 2016: 202, author's translation).

superficial sketches of their inherent logics.[34] Their inner workings are largely obscure to outsiders and even to the engineers who programmed them.

There is another aspect that feeds into the content moderation systems of social media platforms: regulation through legal codes and through the platform's own policies. Though platform companies are largely autonomous in their handling of content moderation (cf. Ciepley 2013: 141; De Gregorio 2020: 7), there are some codified rules that guide these systems. First, there is legislation that generally regulates public communication and also applies to social media platforms. In many countries, communication processes are protected through the freedom of expression. On social media platforms, this expressive liberty is mostly ensured via legislation that exempts hosting providers from liability from third-party content so that these providers are not motivated to drastically censure or moderate the communication processes.[35] But there usually are some constraints in criminal law that regulate what kinds of publicly expressed contents entail sanctions. This legislation applies in both non-digital and digital communication, and it has been the primary means by which nation states regulate digital communication (cf. Pereira/Keller 2022: 2659).[36]

---

34 Due to the complexity, there are multiple points of potential error, many of which have been described by Bianca Prietl (2019). One can assume, for example, that errors may occur when designing the several layers involved in downsizing the content pool, as well as in the specific codes of each layer. Secondly, the input data may be skewed, which – due to the path dependency of learning – would affect all future operations on these data (such as adjustments in categorisation and clustering), as well as how the system self-optimises and consequently how the output itself is being produced. Thirdly, a self-learning algorithm always errs in categorisation, clustering or predicting behaviour, and the standards and ratios of error or success that are applied here form the path by which these systems are developed further. Fourthly, users might misinterpret or manipulate the outputs and thus skew the data on which the system bases its optimisation. Lastly, the systems are constantly re-evaluated in their context of use and optimised by coders. This can result in self-enforcing feedback loops and feedback asymmetries in the platform architectures.

35 The most influential laws in this regard are the US Communication Decency Act (specifically Section 230, CDA) and the Digital Millennium Copyright Act (DMCA), as well as the EU's e-Commerce Directive (cf. De Gregorio 2020; Helberger 2020).

36 Such laws differ from country to country. Some states have anti-hate speech laws like the Danish *racismeparagraffen* (§ 266b of the penal code), and some states aim to prevent misinformation, such as France, with its *loi contra la manipulation de l'information* (LOI n° 2018–1202 du 22 décembre 2018). Infringements of these laws can lead to criminal prosecution. In the international, complex and fragmented space of digital com-

Though many states have authorities that browse the Web for such infringements,[37] the task of finding and sanctioning unlawful contents is mostly performed by the platforms' own content moderation systems. Their terms of service generally prohibit the posting of unlawful contents. On this basis, platforms moderate and filter the posts – this usually takes place after they have been posted. When users see posts they deem illegal (or against community rules), they can notify the platform, whose content moderators and algorithmic systems review the contents and handle them accordingly. Such "[f]lagging is now widespread across social media platforms, and has settled in as a norm in the logic of the social media interface, alongside favoriting and reposting" (Gillespie 2018c: 267). In addition, parts of the platforms' algorithmic systems have often been trained explicitly to filter posts with unlawful contents. Copyright laws, for example, can thus be enforced even before a post is uploaded.[38] When platforms have identified infringements, they can, on the one hand, report these infringements to state authorities, and they can, on the other hand, sanction the offenders on their own platforms through deletion of the offending posts or, at least, by not featuring these posts in other users' newsfeeds.

Moreover, the range of laws that seek to install procedural safeguards directly in the platforms' moderation procedures is increasing (cf. De Gregorio 2020: 10–11; Helberger 2020). For example, the EU's Digital Services Act (DSA) requires platforms to remove specific kinds of contents, to provide complaint mechanisms for the users whose contents have been removed and to disclose

munication, of course, not all users interact under the constraints of the same legal code. What law applies to what user is determined based on the country code in the IP address, but also on other factors. Facebook, for example, was criticised in early 2022 by German NGOs for distinguishing private "consumers" from "all other cases". The former users can sue in German courts whereas the latter users are referred to the laws of Ireland, where the Facebook servers are stationed (cf. Klausa 2022). The diversity of applicable codes of law on social media platforms is countered by the uniform rules that are imposed by the platform companies themselves. They often implement the strictest regulations for all their users – even those who are located where the laws are less strict – because using too many different frameworks is usually costlier than using fewer but stricter frameworks.

37 The German *Medienanstalten*, for instance, have developed an algorithmic system that searches for crimes against human dignity, racism, pornography and the endangering of youths, amongst others (cf. Greis 2023).

38 German copyright law, for instance, compels platforms to use upload filters to scan posts for copyright violations before they go online (cf. Hurtz 2021).

information about their activities, for example, in their terms of service and in regular reports (cf. Klausa 2023: 64). Similar procedural safeguards are included in the EU's Regulation on Terrorist Content (cf. De Gregorio 2020: 10), the EU's Directive on Copyright in the Digital Single Market ("Copyright Directive") (cf. ibid.: 9–10) and the German Netzwerkdurchsetzungsgesetz ("NetzDG") (cf. Klausa 2023: 64). While supporting the transparency and accountability of platform behaviour, this kind of legislation usually does not interfere in the actual content moderation techniques of the platforms.[39]

Platforms additionally have "the right to establish and enforce rules within [their] jurisdiction *beyond* those of the laws of the land" (Ciepley 2013: 141, orig. emph.). Versions of these rules are publicly available in documents like the platforms' community guidelines and their terms of service. The community guidelines specify how users are expected to behave on the platform, what to do and what not to do. The terms of service constitute the contract through which the parties agree to follow and abide by a specific set of rules, including

39 However, a special case of platform regulation is presented by the German Medienstaatsvertrag ("IMT"). Like the above-mentioned laws, the IMT states what contents are prohibited and ensures that users have the possibility of complaining against and counteract platform decisions. Moreover, it provides a "positive discrimination of certain content of value" (Klausa 2023: 64): "by imposing non-discrimination rules as well as the bias towards certain content providers, the German legislator has started to introduce traditional journalistic value sets into the new world of digital media. […] Cautious to avoid prescribing rules and values as such […] they promote legacy actors which already embody said norms. This approach is a difficult balancing act: On the one hand, the state is careful not to interfere with substantive moderation by stipulating mainly procedural aspects like non-discrimination, transparency and alike. On the other hand, it demands a positive bias towards certain content of value without giving the platforms any leeway to choose the respective content themselves" (ibid.: 63). Thus, the IMT supports journalistic standards and actors on the platforms without the state's interfering with substantive moderation (cf. ibid.: 63). This measure thus realises a demand formulated by Habermas, who insists on "reviv[ing] the normative standards guaranteed by journalistic gatekeepers now undermined by the business models of platform corporations" (Thiel 2023: 74; cf. Habermas 2022: 151). It is reasonable to suspect an increasing amount of regulations that actively intervene in digital moderation systems, such as those of social media platforms, in the nearer future. But regulating these systems – ideally in a democratic manner – necessitates that there is a minimal understanding of the platforms and their logics among legislators and throughout society (cf. Coyle/Weller 2020). These processes will be further complicated by the platforms' reluctance to be regulated (cf. Culpepper/Thelen 2020: 301; De Gregorio 2020: 8) or even to disclose the procedures of their systems.

those that are codified in the community guidelines.[40] The content policies, specifically, are rather dynamic: they take up the task of "anticipating violating content" while also "reacting to unprecedented forms of content violation and operationalising new regulations" (Gillespie 2018c: 264).[41]

The platform companies' motivations to establish their specific sets of rules seem to vary. First and foremost, they all have three main incentives: to keep users on their platforms and thus keep generating profit from users (cf. De Gregorio 2020: 2; Gillespie 2018a: 5), "to protect their services from becoming unusable" (Douek 2021: 768) and to keep the platform company out of legal and societal trouble (cf. Stanley-Becker/Dwoskin 2020). But there seem to be further factors that motivate these companies to impose their specific sets of rules in specific manners, including: a sense of responsibility (cf. Roose 2020a); an intrinsic motive to "prevent […] their digital spaces [from] turn[ing] into hostile environments for users due to the spread[,] for example, of incitement to hatred" (De Gregorio 2020: 2); an intrinsic motive to deliver accurate information (cf. Newton 2020f; Roose 2020a); the aim to stay politically neutral (cf. Newton 2020a); the aim to avoid politics on the platform (cf. Lima 2021a); the aim to keep the regard and trust of users (cf. Culpepper/Thelen 2020: 297–298; Tarnoff/Weigel 2020: 151); and the aim to protect their brand (cf. Newton 2020f). So, platforms usually must keep different groups within society content through their policies and moderation efforts; they are "stuck […] halfway between users with different values, halfway between policymakers and the people they seek to regulate, halfway between a conduit and a curator, and halfway between the pressure of internal aims and external demands" (Gillespie 2018c: 272). Accordingly, their rules "are written to encompass multiple interests, not just individual speech rights" (Douek 2021: 763).

Again, the platforms enforce these rules through their content moderation systems, and they have different means of doing so that reflect different levels of escalation.

40 However, De Gregorio (2020: 8) argues convincingly that this kind of contract "leads users in[to] a status of subjection vis-à-vis online platforms", rather than serving as a contract between equal parties.

41 These sets of rules have thus far been largely similar across all the major platforms (cf. Gillespie 2018c: 263–264).

> The basic idea here is to allow for a maximum of political speech, and to answer the most problematic speech with more speech, in the form of labels. The platforms have offered no positive conception of what political speech should be or do there. Instead, they police it as beat cops, running off the worst posts while writing speeding tickets for lesser offenses. The idea rests upon a foundational belief that both parties are good-faith actors when it comes to political speech, all available evidence to the contrary. (Newton 2020b)

So, there are different levels of sanctioning, ranging from adding labels and warnings to posts (cf. Newton 2020e; Silverman 2020), to demoting contents in recommender systems, to preventing content sharing and blocking posts from the view of other users (cf. Ananny 2020: 363–364; Narayanan 2023; Roose 2020a, to deleting posts and banning users from platforms (cf. Roose 2020a; Guardian 2021). Arranged in this order, the sanctions escalate regarding both their level of restriction and their effectivity in reducing visibility. While a label might only lead to a 10% decrease in sharing (cf. Silverman 2020), "a 20% reduction [of visibility through recommender systems] causes [a post's] reach to drop *tenfold*, and the content only reaches the poster's immediate network" (Narayanan 2023, orig. emph.); a platform's ban prevents future postings altogether.[42] These various measures have been combined with other strategies – some platforms provide stable links to established news pages in prominent places, and some platforms add information banners in reaction to coordinated campaigns of rule infringement that seek to influence, with varying success, political events such as US elections (cf. Newton 2020e, 2020f; The Election Integrity Partnership 2021; Wakabayashi 2020).[43]

---

42 This is especially the case if "our policies don't allow people to come back", as was proclaimed in 2021 by the CFO of the website formerly known as Twitter (Guardian 2021).

43 However, content moderation "will always be a matter of probability[, which] makes it unrealistic to expect rules to be applied correctly in every case" (Douek 2021: 765–766). This is because, as I described above, what is happening within a content moderation system is emergent from both the individuals that interact on the platforms and the contextual influences that interfere with each other while structuring the communication processes (cf. Rieder/Hofmann 2020: 8). Moreover, the scale of data and posts that social media platforms process makes it hard for platforms to keep up with assessing and sanctioning contents (cf. Gillespie 2018c: 265). Accordingly, they have to formulate their rules for regulating user behaviour "with awareness of the error rates inherent in enforcing any rule at the truly staggering scale of major platforms" (Douek 2021: 763).

However, the platforms can enforce their rules mostly at their own discretion. They do this "by virtue of the control over their digital spaces […] without the need to rely on a public mechanism such as a judicial order" (De Gregorio 2020: 7).[44] This right has, for a long time, been "relatively uncontroversial" and largely "unencumbered by the constraints of public law" (Douek 2021: 768), though, increasingly, "[p]arliaments and legislatures in all liberal democracies have the regulation of the digital public sphere on their agendas" (Chambers 2023: 67).[45] Nonetheless, the actual content moderation does not necessarily mirror the rule sets codified in the documents, but might rely on internal guidelines that are not disclosed to the users (cf. De Gregorio 2020: 7). In fact, the actual enforcement of platform codes often seems to depend on the policies and business model of the platform, as has been observed in several instances (cf. Newton 2020d; Schröder 2020; Stanley-Becker/Dwoskin 2020). Gillespie thus sees these rules as "statements of both policy and principle – struggled over by the platform operators at some moments and ignored at others, deployed when they are helpful and sidestepped when they are constraining [to the platform company]" (2018c: 264).

In sum, the platforms provide their users with personalised feeds that are based on similarity metrics and the individual user's behaviour and that are additionally guided by legal requirements and self-imposed platform rules. Nonetheless, the platforms provide different possibilities for users to tweak

---

The platforms therefore "are now firmly in the business of balancing societal interests and choosing between error costs on a systemic basis" (ibid.: 762).

44 However, some of the major platforms have established councils to supervise these decisions. For instance, Meta has established the Facebook Oversight Board, X (the website formerly known as Twitter) has the Trust and Safety Council and TikTok has the Content Advisory Council. But still, the companies are only honour-bound – not legally required – to accept the decisions of these councils (cf. Newton 2020c).

45 In particular, the private power of platform companies that allows them to moderate communication at their own discretion "sits uneasily with the primary responsibility and ultimate obligation of states to protect human rights and fundamental freedoms in the digital environment" (Kettemann/Tiedeke 2020). Several legislative bodies have therefore begun to pass legislation that specifically targets social media platforms and their content moderation techniques – an approach to content regulation that is known as the "new school" of regulation (cf. Balkin 2018; Klausa 2023; Pereira/Keller 2022: 2670). It makes use of techniques that "regulate speech through control over digital networks and auxiliary services like search engines, payment systems, and advertisers; instead of focusing directly on publishers and speakers, they are aimed at the owners of digital infrastructure" (ibid.: 4).

their personalised feeds. The degree and form of the influence of these tweaking possibilities on the feed varies according to platform. On most platforms, users can actively follow, flag or even block other profiles and groups; on some platforms,[46] they can subscribe to certain lists, topics or groups that can be used as filters and that – in the case of closed groups – might enable them to see the post and receive the feed notification in the first place; on some platforms,[47] they can generate their own lists according to topic or importance to the profile owner; on some platforms,[48] users can hide certain posts from their feed; on some platforms,[49] they can adjust their settings to prioritise certain profiles, turn on notifications for every new post on a profile or temporarily or indefinitely exclude certain profiles from their feed; on some platforms,[50] they can reset or delete the feedback that the algorithmic system bases its recommendations on, or even only parts of it; moreover, on most of the dominant platforms,[51] they can use apps that are external to the social media platform in order to adjust their feeds. On a few platforms (TikTok being the most prominent example), the feeds are generated almost purely by algorithms. "Merely by watching some videos, and without having to follow or friend anyone, you can quickly train TikTok on what you like" (Wei 2020). These platforms' recommendation systems put more weight on actual user behaviour – such as viewing time, scrolling speed and one-click reactions – than on subscriptions or profile settings. The less influence the users have over their feeds, the more untransparent the feed and its production appear.

The lack of transparency of algorithmic systems is one of the central problems for researchers who try to evaluate the effects of algorithmic systems or to develop ideas on how to judicially govern them.[52] This lack of transparency

46 These include Facebook, TikTok and the website formerly known as Twitter (cf. Freeman 2011; TikTok 2020a).

47 These include Facebook, Instagram and the website formerly known as Twitter (cf. Facebook Help Centre 2022c; Mudit 2022; Freeman 2011).

48 These include YouTube, Instagram and TikTok (cf. Mudit 2022; TikTok 2020a; Writtenhouse 2022).

49 These include Facebook, YouTube, Instagram and TikTok (cf. Facebook Help Centre 2022c; Mudit 2022; TikTok 2020a; Writtenhouse 2022).

50 These include Instagram and YouTube (cf. Writtenhouse 2022; Mudit 2022).

51 These include Facebook, Instagram, TikTok and the website formerly known as Twitter.

52 The same circumstances make digital platforms inscrutable to their users. In fact, a "platform's actual processes may be very different from how they are understood by their Users, who may form mental images of those processes based on their own individual interactions or on how the platform has represented itself to them" (Bratton

has multiple causes. First of all, the companies behind algorithmic systems do not usually share the details of the program codes with the public, since these codes are treated as trade secrets (cf. Bennett/Lyon 2019: 9; Bennett/Oduro-Marfo 2020: 49).[53] But even if the codes were publicly accessible, they would still remain highly obscure. Even apart from machine learning, algorithms are hard to interpret.

> [T]he presence of different programming languages and execution environments adds further complications, and so do the many subsystems and modules that concrete programmes often draw on. Algorithmic decision procedures "may not happen all in one place" (Dourish, 2016, p. 4) but can be distributed over many different locations in a large programme or computer network. […] The result is a continuously evolving metastable arrangement. (Rieder/Hofmann 2020: 7)

Moreover, in order to wrangle the vast amounts of data, the systems need to be as efficient as possible. They can achieve this efficiency by processing the data "in various steps, strategically arranged to make it fast and to limit the amount of computing resources required" (Lada et al. 2021). One such arrangement is what Gilbert Simondon calls "concretization" (2012) and "what technologists themselves call 'elegance'" (Feenberg 2017: 72). "In contrast to a design in which each structure is restricted to a single function, an elegant structure serves many functions at once" (ibid.: 72–73). As devices and algorithmic systems develop, "they are continually redesigned to multiply the functions served by their components" (ibid.). They thus become increasingly efficient, but also increasingly intricate and inscrutable.

2016: 50). Consequently, each user adapts to their own perception of how the platform is working and is responding to their behaviour.

53 Of course, there are considerable similarities between the recommender systems of the major platforms. In fact, "most major platform companies are quite open about discussing their recommendation algorithms at academic and industry conferences, and learn from each other. […] But it turns out that the set of details that matter from a research and engineering perspective are subtly different from those that matter to users and society. And companies seem to have little incentive to be transparent about the algorithm with those stakeholders. […] Besides, companies have shared precious little about the effects of algorithms" (Narayanan 2023). However, when the exact source code of Twitter was leaked in early 2023, there was great consternation (cf. Mac/Conger 2023).

Furthermore, the criteria for selecting one content over another are mostly hidden. "What is unstated is how these criteria are measured, how they are weighed against one another, what other criteria have also been incorporated, and when if ever these criteria are overridden" (Gillespie 2014: 176). They are hidden either deliberately by the platform (ibid.: 192) or in the constant self-optimisation of machine-learning techniques. The reason for this is that the coders do not devise the actual rules for algorithmic ordering, but they set the rules for how the system is supposed to learn to reach a certain aim. The resulting strategies are often so complex that they cannot even be understood post factum. It is difficult to interpret them experimentally, and they cannot be interpreted logically.

Lastly, the algorithmic systems of social media platforms are highly dynamic. On the one hand, data change with every click of every user. For example, "[t]he tweets we post, items we purchase, and likes we hand out all factor into the sum total", and this, in turn, "modif[ies] what becomes visible to us and to our networks" (Bucher 2018: 154). Therefore, every query delivers slightly different results. Systems including machine-learning algorithms are even more internally variable. They "evolve by definition as they encounter new data" (Cotter/Reisdorf 2020: 747), because platforms "are not just providing information to users but also users to their algorithms. And algorithms are made and remade in every instance of their use because every click, every query, changes the tool incrementally" (Gillespie 2014: 173).

On the other hand, platform companies often adjust their policies and services in order to keep their clients and users, to gain new ones and to generally stay ahead of their competition. Moreover, they develop their systems to make the systems more "elegant" and efficient (cf. Feenberg 2017: 72), and the aim of doing this is to handle the ever-growing amounts of data without unnecessary delays. Their algorithmic systems are constantly "subject to examination, intervention, and refinement" (Healy 2015: 193). They can be "easily, instantly, radically, and invisibly changed" (Gillespie 2014: 178). The "iterative design of algorithms and A/B testing of different iterations mean that to some extent algorithms 'never take durable, observable forms'" (Cotter/Reisdorf 2020: 747).[54] At the same time, user habits and demands change, providing another reason for

54 Changes to the code are "inherently framed and shaped by all kinds of decisions, politics, ideology and the materialities of hardware and infrastructure that enact its instruction" (Kitchin 2017: 17–18).

platform companies to adjust their systems – and re-adjust them if the original adjustments backfire (cf. Newton 2022b). Moreover, policy adjustments are visible in preparation for, during and in the aftermath of important events, such as elections, pandemics, terrorist attacks or other crises.[55]

In sum, there are some general logics that are incremental to social media platforms. First of all, they rely on vast amounts of data. The platforms standardise these data for the purpose of processing, and they process the data for diverse reasons, including the moderation of interaction on the platform. Central tools of content moderation include algorithmic recommender systems that, based on similarity metrics and statistics about the past behaviour of individual users, predict what contents will generate engagement by what users. Moreover, the platforms are guided in their moderation procedures by legal requirements and by their own, self-imposed rules, which are formulated in community guidelines and terms of service. But beyond these general logics and guidelines, algorithmic content moderation systems are largely untransparent and highly variable. Even the self-imposed platform rules do not allow reliable conclusions to be drawn about the details of platform procedures. As a result, it is impossible to generalise the effects of platforms' algorithmic systems in a scientifically sound manner or to pretend that detailed analyses about such effects have a long half-life. My own analyses must therefore be limited to what is actually known about these systems and to the more stable logics embedded in the systems.

4.2 The Relevance of Social Media Platforms for Deliberative Systems

The notion that social media platforms are relevant for democracy is as old as the oldest platforms, and I have repeatedly alluded to their relevance for deliberative systems already. These platforms are central communication and

55 One issue that usually appears in these contexts and that has gained momentum in the public debate concerning social media platforms is misinformation. It has been discussed in the context of the US presidential elections, and specifically in the context of the COVID-19 pandemic and the vaccination campaigns. Platforms have introduced different systems to prevent misinformation, including a system to reduce the distribution of all posts that are from people who routinely share misinformation (cf. Wagner 2021), as well as flagging systems and fact-checking by human content moderators (cf. Ananny 2018; Gillespie 2018c; Marres 2018).

information-gathering infrastructures. Communication on these platforms mostly occurs in more informal manners, but information on this communication is transmitted to all kinds of spheres in the system. These transmissions thereby amplify informal communication processes between individual citizens and make these processes relevant to a wide variety of sites within the entire system. Moreover, the platforms structure the communication processes that take place on them in accordance with their own logics and rules. Due to their oligopolistic tendencies, few platforms gain large amounts of resources and influence in different parts of deliberative systems through their structuring and through the transmissions into the wider system.

4.2.1 An Oligopoly of Platforms

Social media platforms are highly relevant for individuals and societies under the conditions of the digital constellation. Since digital platforms connect third parties with each other, they "*set the stage for actions* to unfold through ordered emergence" (Bratton 2016: 47, orig. emph.). Large parts of the communication processes that are thus structured occur on only a small number of social media platforms. This is because platform companies have developed "an immanent, quasi-natural tendency to build monopolies, respectively oligopolies" (Vogl 2021a: 10, author's translation; similarly, Jungherr et al. 2020: 25; Zuboff 2022: 3). This process is largely caused by the effects these platforms have on their users, such as the lock-in effect, the network effect and the log-in effect.[56]

The "lock-in" effect (Arthur 2014), or the "generative entrenchment" effect (Bratton 2016: 47), describes a situation in which users have gotten used to a certain platform's interfaces, standards and services to such a degree that their staying on this platform is easier and needs less effort than switching to another platform with other standards (cf. ibid.). For example, most social media platforms provide specific editing programmes for the formats they support. Instagram has its own photo editing programmes while TikTok allows video editing in its app. Getting used to working with these editing programmes that only work with the respective platform – or at least do so considerably better – may induce the users to stay on the respective platform.

Moreover, social media platforms have a network effect that causes new users to sign up for them and established users to continue using them (cf.

56 Moreover, the process has been "systematically facilitated" by "legal institutions, including both entitlements and regulatory institutions" (Cohen 2017: 133).

Jungherr et al. 2020: 24; Vogl 2021a: 10). If "everyone I know" interacts on a certain platform, I do not want to miss out. Or, put more generally, the more users there are on a platform, the more attractive it gets, and as it gets more attractive, more users will join. Additionally, large user numbers help these platforms improve their services – for example, the precision of their recommender systems – which makes them even more attractive for users and advertisers (cf. Muldoon 2022: 30), which in turn draws in even more of both. Consequently, user numbers of already big networks exponentiate, which contributes to "a tendency for mono- or oligopolies to emerge" (Jungherr et al. 2020: 25).

Exponential effects similar to the network effect can be witnessed with regard to the "log-in" effect, which refers to the user's time spent on a platform. "The use of a platform for any user increases exponentially with the number of all available users on the platform. Linear user growth on a platform thus still leads to exponential growth in the use that users are expected to make of the platform" (ibid.: 24–25). Encouraging users to spend as much time as possible on the platform is one of the central concerns of platforms' architectures and moderation systems.

Hence, platform companies aim to make users join the platform (the network effect), stay on it (lock-in effect) and use it (log-in effect).[57] Due to their network, lock-in and log-in effects, large platforms tend to become even larger; and, ultimately, a few platforms dominate large shares of the various resources in an oligopoly of platforms. Next to money and user numbers, these resources include users' senses of community, users' time spent and engagement on the platform, data, advantages from machine learning and, ultimately, money – all of which are inherently connected. That is, as users spend more time on a platform and interact there with an increasing number of other users, their sense of connection becomes greater. "The largest of the global networks exercise a near monopoly over this resource" (Muldoon 2022: 50). Of course, there are "avenues to pursue community in real life and on other networks online, but the dominant position of a few tech companies cannot be denied" (ibid.). The sense of community further intensifies platform effects and oligopolistic tendencies. As a result, a few large platforms have large shares in users' time and engagement. This means that only a few platforms can generate comprehensive datasets concerning user behaviour; the most detailed user profiles are generated by only a few platforms, and these platforms secure their exclusive

57 The reverse effects are also possible for platforms that suddenly seem unattractive, as has recently been demonstrated by Twitter/X (cf. Statista 2024), among others.

access to these data by erecting technological barriers around them (cf. Pistor 2020: 123).[58] Companies that own several platforms can combine the respective databases, thus deepening the tendency towards a data oligopoly for each of their platforms. In the wake of data oligopoly, an oligopolistic tendency regarding the advantages of machine learning follows, since

> what success [machine learning] is liable to achieve will tend to accumulate over time through the growth of the datasets and past trajectories of machinic interaction with them. This cumulative advantage creates an oligopolic tendency because there are few players with the computational or human resources to fully take advantage of these opportunities. (Carrigan/Fatsis 2021: 38–39)

The user number oligopoly, the data oligopoly and the machine-learning oligopoly support further oligopolistic tendencies regarding the market share in ad placement. Where many users are constantly logged in, there are many potential views, clicks and buys, which is the ultimate aim of placing an advertisement. Larger datasets, more accurate user profiles and more potent machine learning enable more effective ad targeting to specific user groups, making the larger platforms even more attractive to advertisers (cf. Gillespie 2018a: 19; Vogl 2021a: 15). In other words, oligopolistic platforms have control over massive amounts of these assets, and it is only this control that can be effectively monetised:

> The value of data is not revealed by the price others are willing to pay for data points or even the sum of all these data points, but by the processing and analytical capacity of the data controller. This capacity depends on scale; that is, on the volume of data already captured and on future access to data. (Pistor 2020: 112)

58 Simultaneously, platform users are practically forced to provide their data to the dominant platforms, since "individual consent is tied to services that most data producers want and for which there are increasingly few, if any, viable alternatives. Individual consent by dispersed data producers who are faced with a take-it-or-leave-it option is not very meaningful, especially in a situation where they face oligopolies or even monopolies that control access to digital platforms" (Pistor 2020: 110). The consent that users give to platforms for collecting and using their data is therefore not given in an entirely voluntary manner.

The clients buying the advertisement spaces that are predicted to be most effective, meanwhile, "have no way of discovering market prices through repeat transactions" (ibid.). This is because the specific effectivity depends on numerous variables, including the time of day and various characteristics of the users who are targeted by the ad. "Every data application is unique and allows the seller to set prices without a market test" (ibid.). Critical political economists, moreover, emphasise "(1) the inherent accumulative tendency of capital and corporate ownership and its subsequent effects on the distribution of power and (2) the precarious and exploitative nature of cultural and (immaterial) labor of both producers and end-users" (Nieborg/Poell 2018: 4279). Consequently, "new digital hierarchies [emerge] between those who exercise control and decision-making power over platform architecture, protocols and rules and those who must submit to these conditions" (Muldoon 2022: 53). The larger the platform, the bigger its sphere of influence. The "globespanning nature of these firms is leading to a transformation of existing international hierarchies and the creation of new forms of digital colonialism" (Srnicek 2020: 89–90).

This even translates to the physical world and the material infrastructures of these platforms. The existence and characteristics of the material infrastructures[59] depend on the geographical, economic and political conditions of the respective region, as well as on who owns the infrastructures and the rights to build them. If the oligopolistic platforms invest their monetary resources in these infrastructures, they can further increase their data resources. In the Global South, for instance,

> Big Tech CEOs act as envoys, meeting heads of state and offering shiny deals of free infrastructure and services in exchange for the collection of data and control over the system. This includes cloud services to government agencies, building broadband cables and even using high-altitude balloons to create an aerial wireless network connecting remote areas of the country. Facebook's notorious "Free Basics" offers a bare-bones version of the internet to countries in the Global South who receive free access to a limited number of data-light websites and services. (Muldoon 2022: 133)

59 Digital communication depends on material infrastructures such as wires, servers, internet junctions and satellites and other transmitters. These newer infrastructures, in turn, "rely on the older infrastructures of water, energy, and human transit, as a nervous system relies on a circulatory system" (Stine/Volmar 2021: 14).

The ownership of these central infrastructures determines who controls them, what services can be accessed, what laws apply, what filters are in place, who collects the data, who profits from the services and in what ways they can profit.

Accordingly, the dominant platforms and their corporate owners wield an "awesome power" (Newton 2022a). They "have the power to *frame the social world, name* it and in the process, categorize and *order* it through their algorithms… changing how the social world is organised […] and changing how we *imagine* the social world" (Couldry 2021, orig. emph.). Consequently, they change the conditions under which democratic decisions are being made.

The implications of the oligopolistic social media platforms regarding democracy are rather Janus-faced:

> Social media is the site of much hatred and delusion. It is also the soil where social movements can take root. Hashtags enable individual incidents of injustice to go viral, revealing systemic patterns. Yet that virality is made possible by monopoly – it requires one Facebook, one YouTube, one Twitter. And these large concentrations of unaccountable private power endanger the basic premise of democracy – the idea, often invoked but rarely attempted, that the whole of the people should determine how society is run. (Weigel et al. 2018: 14)

In other words, as the platform becomes larger, so does the group of users with the potential to interact with each other, independently of time and space, and this increases the number of potential connections between different users, sites and site clusters. The dominance of a few platforms provides "important affordances for bottom-up organizing, collective creativity, and crowd-sourced, democratic action" (Cohen 2017: 153), since many individuals meet in few places. On the other hand, such resources can be used for various purposes that are problematic for democracy. As the platform becomes larger, the amount of communication that is moderated and structured by one single actor becomes larger, and this means that the influence of this not democratically legitimated actor becomes larger. Thus, the "particular configurations that those technologies have assumed within the political economy of informational capitalism also make them sites of extraordinary manipulability, creating new risks [for] the human project of democratic, inclusive, sustainable coexistence" (ibid.). Or, put more succinctly, "[o]ligopoly in the economic realm shades into oligarchy in the societal realm" (Zuboff

2022: 1). The control over the data harvested from their users gives platforms "unchecked power to impose their economic or political preferences on the data producers" (Pistor 2020: 110): "While not every data point, in isolation, that is captured raises human rights concerns, amassing enough information to constitute digital identities of individual data producers and using that information to influence their behavior to the benefit of the data controller does" (ibid.). The consequence is a "democratic deficit refer[ring] to the fact that private companies make the choices that set norms and directly influence the behavior of billions of users" (Haggart/Keller 2021: 2). Platform oligopoly thus simultaneously represents chances for and threats to deliberative systems.

The oligopolistic tendencies resulting from platform effects explain in part why some platforms obviously arrive at and remain in the centre of public, political and academic attention. The published studies that are relevant to my analyses in chapters 5 and 6 will largely pertain to the platforms Facebook, Instagram, YouTube, Twitter/X, Reddit and TikTok. All of these platforms adhere to the logics I have described above. Moreover, they are oligopolistic in the sense I just illustrated, as they rank among the social media with the highest user numbers worldwide (cf. Statista 2023) and consequently have greater access to all the other resources I referred to. These platforms have different aims, interfaces and features; they focus on different forms of content and communication. But incidentally, users find ways to introduce political topics and communication about political opinions on all of them – though arguably to different degrees. All of these platforms are therefore potentially part of deliberative systems; in fact, they are highly interwoven with these systems.

4.2.2 Social Media Platforms' Interwovenness with Deliberative Systems

There are several aspects to how and why social media platforms are practically woven into the fabric of deliberative systems. On the one hand, this is due to the platform's centrality as a communication infrastructure. Platforms serve as a central infrastructure for communication, social connections and communities.[60] The platform effects described above lead to the tendency of very many people to use very few platforms. This, in turn, results in many users' experiencing a "fear of missing out" if they do not join and regularly use the platforms being used by so many others from their immediate and their not so

60 Cf. Bratton (2016: 46), Carrigan/ Fatsis (2021: 35–36), Couldry/Hepp (2017: 7), Dolata/ Schrape (14).

immediate social environment (cf. Bloom et al. 2019: 1251). By provoking such effects, "platforms can make users feel like they could not live without the social […] benefits that accompany access to their network" (Muldoon 2022: 30). While "there is a significant minority who don't use social networks and seem to get along just fine" (ibid.: 50), many people use social media platforms primarily to access community. Dependence on these digital connections can vary widely, ranging from using them for mere entertainment to being dependent on them, due to professional reasons, for one's livelihood. The latter applies specifically to professional political actors who are dependent on information about and communication with communities and other political actors. The increase in the proportion of communication that is being performed via social media platforms also means that political communication is increasingly performed on these platforms. This in turn implies an increasing relevance of these platforms for deliberative systems.

On the other hand, the influence of social media platforms is attributed to their logics and modi operandi, and especially to their content moderation practices. Social media platforms were among the first digital spaces where many-to-many communication could be realised, and content moderation is their main way of handling the vast amounts of communication. But

> [p]latforms don't just mediate public discourse, they constitute it. They are designed to invite and shape participation, toward particular ends. This includes what kind of participation they invite and encourage; what gets displayed first or most prominently; how navigation tools direct the movement of users and content; how revenue models impose price mechanisms; and how they organize information through algorithmic sorting, privileging some content over others, in opaque ways. And it includes what is not permitted, and how and why that prohibition is enforced […]. (Gillespie 2018c: 257)

Hence, platforms shape participation and structure communication processes. They "insert themselves into the democratic process by changing who can communicate with whom and how, and in doing so, who can participate in the political process" (Jungherr et al. 2020: 225). They define the algorithmic logics and design the interfaces that shape the dynamics of reception and communication (Schrape 2021a: 160).[61] By moderating communication, platforms

61 I will elaborate on this in chapter 5.

are also "curating" the information users are presented with and have access to. "The curation of *political* information through secret algorithms gives the social media platforms enormous power over our worldviews, and extraordinary abilities to modify our political beliefs and behaviours" (Bennett/Oduro-Marfo 2020: 49, orig. emph.).

Both these aspects – their centrality as a communication infrastructure and their shaping and intervening in communication processes – make social media platforms relevant for and influential in deliberative systems that operate under the conditions of the digital constellation. They influence not only the respective communication processes but also the preconditions and the context, as well as the consequences and follow-ups of these processes in the digital and the non-digital realm. Thus, the mediated communication on platforms "is interwoven with our face-to-face communication in manifold ways" (cf. Couldry/Hepp 2017: 16).

Almost all of the sites that are mediated through social media platforms are informal in nature, but many of these informal sites are highly relevant to and connected with sites of semi-formal and highly formal communication. Informal sites can include sites that are not explicitly political in nature, such as "online communities dedicated to lifestyle issues such as personal finance, parenting/childcare, popular culture, sports, and hobbies" (Wright et al. 2018: 74–75). They "foster a connection between the personal and political and can potentially help bridge the gap between the everyday lives of participants and formal politics" (ibid.: 75). In such spaces, it is also "more likely that people will inadvertently […] come across people with divergent political views as social boundaries appear to be weakened online" (ibid.: 83). Moreover, deliberative systems attribute to such informal spaces

> the potential to enable more malleable and innovative forms of political expression, such as storytelling and the use of diverse audio and visual tools […]. And since audiences can co-produce scripts, question meanings, and reconfigure debates, there is potentially greater political improvisation. These features suggest that social media could help facilitate broad dialogic communication […], as well as open up space for new voices to enter debates and potentially form alternate alliances that cut across, or bridge, opposing positions. (Hendriks et al. 2016: 1105–1106)[62]

62 In chapter 6.1, I will come back to the implications that a widened range of admissible forms of expressions on social media platforms has for deliberative systems.

Social media platforms should therefore be considered in any analysis of deliberative systems within the digital constellation, if only because of their relevance for informal political talk.

But social media platforms are also relevant for sites of semi-formal communication, such as spaces of traditional media (cf. Conover/Searing 2005: 270).[63] The platforms' importance to a majority of individuals throughout society renders them a central infrastructure for both traditional media and individual journalists (cf. Newton 2023). But in addition to distributing their contents on social media platforms, journalists also use these platforms to collect information on and assess the societal relevance of current events. This is because the users themselves have become part of the flow of political communication on social media platforms by individually posting and sharing information, comments or public exchanges – whether on their pages or their (semi-)public profiles on social media platforms – and by collectively exerting influence on moderation algorithms via their engagement with news items (cf. Jungherr et al. 2020: 11). The rise of social media platforms has consequently "increase[d] the number of actors injecting and amplifying items in the flow of political information […] and 'tether[ed]' political elites, journalists, and audience much more closely to one another" (ibid.).

Furthermore, social media platforms facilitate the initiation, organisation and momentum of social movements – which are likewise sites of semi-formal communication (cf. Fleuß et al. 2018: 17). With their billions of users, the platforms provide a central infrastructure for all these processes (cf. Seeliger/Sevignani 2021c: 30), while simultaneously structuring these processes according to their own logics. Individuals and organisations can make their cases to a potentially broad audience, they can organise events in real-time and issues can spread geographically and socially very far in very short time spans.[64] But the degree to which a platform supports the spread of politics in general, and

63 In chapter 3.2, I already addressed some more general aspects of how the digital transformation affects traditional media.

64 There even have developed some purely digital variants of social activism, campaigns and fundraisers. For example, the "ALS Ice Bucket Challenge" has traversed several different platforms since 2014. In this specific challenge, users were dared to empty an ice bucket over their heads and to donate to several organisations concerned with amyotrophic lateral sclerosis (ALS). The campaign received world-wide participation and donations. On the one hand, challenges resemble non-digital campaigns, as the respective videos are very similar and convey the same simple messages. On the other hand, they could hardly be realised without social media platforms that have such a

individual political issues more specifically, depends on its content policies, it moderation systems and the applicable national law applicable. While "people around the world continue [to] view social networks as critical venues for protest and debate", one of the "biggest questions of the next-half decade will be in how many places Facebook, Twitter, YouTube and others can live up to that ideal" (Newton 2021a). Some platforms explicitly do not engage with politics, and others are retreating. TikTok, for example, aims at being purely for entertainment; ostensibly, its "mission is to inspire creativity and bring joy" (cf. TikTok 2020b). The platform designs its interfaces and algorithmic systems accordingly (cf. Wei 2020).[65] Meta – at that time Facebook – used to engage with politics. It explicitly promoted a global community: "an inclusive community that reflects our collective values and common humanity from local to global levels, spanning cultures, nations and regions in a world with few examples of global communities"; "a civically-engaged community in a world where participation in voting sometimes includes less than half our population" (Zuckerberg 2017). After having had some bad experiences with participation in collective movements in the direction of the US Capitol in 2021, however, they announced a reorientation of Facebook and Instagram predominantly towards connecting friends and families (cf. Lima 2021a; Newton 2022b).

Furthermore, social media platforms are connected with the highly formal sites of the empowered spheres of deliberative systems. For example, political actors in empowered spaces can be just as affected by social media platforms as any individual in the digital constellation. But due to their political relevance, the implications of many-to-many communication and moderation are different. Political actors can be monitored by individuals and institutions, their digital communication with the public can be mediated or unmediated, their perception of citizens and citizens' wishes is shaped by the moderation and interfaces of the platforms and their political information environments become fragile due to the dynamics on the platforms (cf. Jungherr et al. 2020: 5). Thus, the impact of social media platforms on the empowered sphere "is pervasive,

high number of participants and a global reach and that take place over such a long stretch of time.

65 However, political communication and the organisation of social movements via the platform cannot even be avoided by TikTok (cf. Cockrell 2019; Dornis 2020; Lorenz et al. 2020; Gilbert 2022; Lorenz 2020b, 2020c; Lorenz et al. 2020; Medina Serrano et al. 2020, 2020; Washington Post 2022).

albeit not necessarily transformative in changing fundamental characteristics in practiced democracy" (ibid.).

Conversely, participants in empowered spaces can use social media platforms to monitor the public. As I described in previous chapters, platforms collect data from interactions taking place on social media platforms, and representative public opinion research is increasingly complemented with, if not supplanted by, analysis of these data (cf. Dormal 2021: 16). "The possibilities to analyze and predict behavior by collecting and combining all digitally available traces in order to 'read' the public have begun to compete with participatory modes of will formation that are traditionally identified with the democratic process", and such techniques' promise of "comprehensive and neutral representation is appealing to decision makers and can change expectations and structures over time" (Thiel 2023: 73).[66] In consequence, data on informal and semi-formal communication processes become increasingly important to the spheres of political decision-making.

Yet if political actors equate social media data with public opinion polls, they interpret these data as mirrors of the users' preferences, which generates several problems. Firstly, as I have already shown, these data are singularised, and their accumulation is not equivalent to public opinion formed by open, ideally deliberative communication.[67] Secondly, these data are not only based on what users explicitly mark as their political opinion but are often gathered from the users' behaviour without their knowledge or intention. Therefore, the data do not reflect what citizens want their representatives to understand as their political opinions and wishes. Thirdly, the demographics of platforms widely differ from those of nation states or most other constituencies. For example, platform users tend to be younger, more female and better educated than the general demographics in their respective countries (cf. Gambo/Özad 2020; Mellon/Prosser 2017; Ribeiro et al. 2020; Wojcik/Hughes 2019). Lastly, data are immensely coloured by the platforms' behaviour toward their users. User interaction is shaped by the designs and affordances of the platforms' services, as well as by the algorithmic systems that are moderating communication. Thus, there is an imbalance between, "*on the one side,* […] *users with 'influence-able' and 'target-able' opinions, tastes, and preferences,* [and], *on the other side,* […] *authoritative data analysts who 'know' the population's fine-grained and ever-changeable preferences and tastes*" (Marres 2018: 437, orig. emph.). Future

66 In chapter 3.2, I described this in more detail.

67 See chapter 3.2.

communication is moderated according to this "knowledge", generating more data and continuing a feedback cycle. The data regarding user interactions are therefore heavily shaped by the platforms' rules and processes. All in all, if political decision-makers are placing weight on the data of social media platforms, they are placing weight on a skewed image of a non-representative part of society, an image that is accumulated from singularised data and heavily coloured by the platforms' own conduct.

In sum, platforms are interwoven with various spheres of deliberative systems. They are central communication and information-gathering infrastructures. Communication mostly occurs in more informal manners, but information is gathered by all kinds of spheres within the system; consequently, the platforms amplify informal communication processes between individual citizens so that these become relevant to a wide variety of sites within the entire system. Simultaneously, they structure the communication processes in accordance with their own platform logics and rules. Due to their oligopolistic tendencies, a few platforms gain large amounts of resources and influence in different parts of deliberative systems.