Democratic Autonomy vs. Algorithms? Limits and opportunities for public reasoning

Sophie Jörg

1. Autonomy and Technology: A glance at a long tradition

Thanks to the ever-growing availability of data and the increasing power of algorithms, political practices and processes are changing. But what are the implications of the increasing use of algorithms for people's political or democratic autonomy? Do algorithms have an *empowering* or *disempowering* effect on us citizens? Drawing on references from the philosophy of technology as well as the theory of democracy, this paper makes the case for a differentiated, functionality-based approach towards the impact of algorithms for public reasoning. By shedding light on the curation, moderation and verification of information by algorithms, it is argued that democratic autonomy – i.e., individuals' jointly reasoned self-rule or public reasoning – is empowered by algorithms if they improve the quality of discourse and utilize the process conditions of public reasoning as benchmarks for moderating content. Whereas algorithms disempower citizens if they are used as automated shortcuts to the presuppositional process of jointly reasoned self-rule of citizens.

The public debate about the social and political impact of aritifical intelligence (AI), machine learning- (ML), and deep learning- (DL) technologies in liberal democracies seems to be overtaking the earlier concern about the replacement of humans by machines. In this context, amongst others, it is discussed how digital infrastructure alters the process of democratic elections and public opinion-forming, the voter turnout as well as the voter behavior, or how DL can affect political communication in general. Against this backdrop, new theories of democracy – such as E-Democracy (e.g., Kneuer 2019),

Data Democracy (e.g., Batarseh/Yang 2020), or Liquid Democracy (e.g., Paulin 2020) – emerged.

Yet, all these issues – raised in political philosophy as well as political science – resonate with the fundamental question of both the humans' and the technology's *agency* and their respective *autonomy*: Do digital technologies in fact expand the scope of human action or do they restrict it? How does this affect democratic practices and their political legitimacy? By contrast, what kind of agency is ascribed to DL-technology, for example?

Although the debate about human autonomy has experienced a renaissance in the wake of technical automation, exploring the relation between technology and autonomy has a long tradition as it has always been a key focus of machine ethics and philosophy of technology (Heßler 2019). Autonomy, here, is essentially negotiated on both levels: The autonomy of humans on the one hand and the autonomy of technology on the other; resulting in a twofold meaning of the concept: human autonomy and technological autonomy or ›autonomous systems‹ (cf. Chiodo 2022).

However, with the recent advances in DL and ML algorithms and their ubiquitous use, the long-standing research on the correlation between technology, autonomy, and automation has gained new momentum. The improvement of complex algorithms inscribed – or rather coded – technology with a higher degree of agency resembling in a decisive caesura regarding human-computer-interaction (HCI). To date, the effects of this caesura on human autonomy have been studied primarily at the individual level, along the lines of the enlightenment concept of self-determination (individual or personal autonomy) (cf. Laitinen/Sahlgren 2021; Sankaran et al. 2020). Whereas the systemic dimension of autonomy in the political sense of self-legislation (political or public autonomy), as coined especially by liberal theorists of democracy, is comparatively omitted – despite the prominent references to the potential threats Tech might pose for democracy (cf. Sætra 2021).

The aim of the paper is to further differentiate the analysis of human autonomy in the context of algorithm-based technology with regard to the democratic theoretical implications of the algorithm-driven shift in HCI. In doing so, the reflexive consideration of individual's personal autonomy shall be extended to the notion of autonomy as self-rule on group level, as Max Weber might say (Weber 2006 [1948]). To this end, human autonomy will be considered not only as an (intrinsic) end in itself, but as a necessary and functional element in the fabric of democratic practice. Against this background, the paper discusses both the drawbacks and merits of algorithms and algorithmic

decision-making in the public sphere for citizens realizing their democratic autonomy.

2. Autonomy and Algorithms: Conceptual references for an analysis

Everyday life is permeated by digital, algorithm-based search engines (e.g., Google, Bing), translation programs (e.g., Google Translator, DeepL), or recommendation systems (e.g., AboutYou, Netflix, Spotify) that serve as decision filters for various purposes. Algorithms thus not only influence our taste in clothes, movies, or music (Goldschmitt and Seaver 2019), what information we receive in our social media news feed (Pentenrieder 2021, 53), they might also determine our credit score or political attitude (Cho et al. 2020). »Algorithms«, according to Verständig et al. (2022, 8), »are invisible and yet ubiquitous and tangible«[1]. So, in some respect they might actually determine our thoughts and actions. But does this ›algorithmization‹ or ›platformization‹ of the public sphere and communication, as recently postulated by Habermas (2022) or Eisenegger (2021), also amount to a dilution of the legitimacy of political decision-making in liberal democracies as it infringes people's democratic autonomy? To further address this question, we will first review the conceptual roots of the two main concepts – Autonomy (2.1) and Algorithms (2.2) – in light of the aforementioned triad of automation, autonomy and democracy.

2.1 Democratic Autonomy as *jointly reasoned self-rule*

During the European Enlightment the philosophical concept of autonomy found its way into Western societies and has been regarded as a core concept of modernity ever since (Thimm/Bächle 2019, 75). As an antonym to external determination, it originates in Greek (›*auto*‹ = *self*; ›*nomos*‹ = *law*) and refers to an individual's self-determination or self-rule through which people can participate as subjects on their own authority (Quante 2013, 47). With Kant, the term was linked in the 18th century to the use of human reason, which affected the discourse about supposedly ›autonomous technology‹ to this day (Heßler 2019, 249). With the prominent saying ›*sapere aude*!‹ (*engl*. ›Have courage to seize your own reason!‹), Kant's philosophy paved the way out of mankind

1 Translated by the author.

from ›self-inflicted immaturity‹. By means of the reason bestowed upon them, humanity could free itself from paternalism and foreign domination. Personal (or individual) autonomy in this context means the self-determination of the human subject through reason. Consequently, autonomous subjects are those who have their own freedom of will and decision by utilizing their ability to think critically and revise their stances respectively (Pohlmann 2010).

In the 20th century, theorists, such as Habermas (1992) and Rawls (1996), further differentiated the notion of autonomy in terms of its political dimension (cf. Thimm/Bächle 2019, 76). Rawls, for example, conceptualized autonomy – similar to ancient Greece – as a political category for the preservation of citizens' freedom. Unlike personal autonomy, this type of autonomy first appears in the political sphere or the political activity of citizens (Weithmann 2011, 327). According to Rawls, it would be realized by citizens »[…] participating in society's public affairs and sharing in its collective self-determination over time« (Ibid.). Habermas, however, argues that this shade of civic or public autonomy should not be considered separately from the private or personal autonomy described at the outset. On the contrary, both concepts are co-original (*gleichursprünglich*), mutually dependent and equally fundamental (Habermas, 2019, pp. 134, 163; Jörg 2023, 4): Public autonomy, in a nutshell, presupposes private autonomy, because it requires a legal system that is legitimate only if it guarantees equal freedoms to its subjects. Private autonomy, on the other hand, requires public autonomy, as the legal regulation of private autonomy is legitimate only if it emerges from a discursive process that guarantees political rights.

In regard to the emphasis on civic self-legislation, Habermas's understanding of autonomy is described in the literature as an »attempt to rethink the Kantian (and Rousseauian) idea of individual freedom through self-governance« (Anderson 2019, 22). The concrete exercise of self-governance, then, takes place within the framework of the public, discursive formation of opinion and will (Habermas 2019, 161). Accordingly, political autonomy can only be achieved by »equal opportunities to participate in processes of opinion- and will-forming«[2] (Habermas 2019, 156). In this respect, the entangled concepts of public and private autonomy are essential for a functioning democracy, as they pinpoint the necessary capacity of humans to (self-)reflect, critically assess, and soundly evaluate solutions regarding complex political, social, or individual conflicts (cf. Richardson 2002).

2 Translated by the author.

These philosophical references, however, inform the conception of the so-called democratic autonomy, which in essence describes the principle of people's jointly reasoned self-rule. In this notion of human autonomy one of the most central – if not the most central – democratic maxim culminates: the primal demand that »we [the citizens] must reason together in order to rule ourselves« (Richardson 2002, 18). Realizing citizens' democratic autonomy, hence, is constitutive for the democratic constitution of political order and its legitimacy. Drawing on the co-originality of public and private autonomy, then, the idea of being »the author of my private life« (*personal autonomy*) is closely tied to the idea of »joint authorship of our common political life« (*public autonomy*) (Lovett/Zuehl 2022, 469). With this merger of individual self-determination and systematic self-rule or self-governance both influential manifestations of autonomy are conceived together and united in the term's core as ›public reasoning about the ends of policy‹ (Richardson 2002). Drawing on Habermas' influential theory of democracy, this public reasoning, in turn, is closely linked to four crucial – albeit demanding – preconditions: The (i) *inclusivity* of all possible topics and participants, (ii) *equal* distribution of *participation* opportunities, the (iii) *sincerity* of all participants, and finally the (iv) *absence* of a *coercive communication structure* (Habermas 2009, 89). Accordingly, democratic autonomy or jointly reasoned self-rule manifests itself in the process of public reasoning; meaning the democratic process of public opinion-forming and the public decision-making based on it.

2.2 Algorithms as mechanism of technical *automation*

Algorithms are often recognized as the technical underpinning for technology's alleged ›agency‹ and ›autonomy‹. In the following, however, it is argued that they equal mechanisms of pure technical automation, which are often misinterpreted as autonomous, i.e., self-determined, action of technology. What may initially appear to be autonomous action in the course of technical processes actually turns out to be nothing more than a complex automation of various computing steps programmed by algorithms in order to perform specifically defined tasks. While exploring the automation of technology by algorithms, possible limits of a ›technological autonomy‹ are briefly pointed out.

However, to understand why exploring algorithms is crucial to be aware of the implications of digital technologies for democratic legitimacy of political decision-making, it is worth taking a glance at the technical aspects of

algorithms: In computer science an algorithm is acknowledged to be a »list set of instructions, used to solve problems or perform tasks, based on the understanding of available alternatives« (IIG 2023). Algorithms are systematic and mechanical methods for solving a well-defined problem, such as efficiently searching or sorting vast amounts of data (Belford/Tucker 2023). Due to detailed »specifications for performing calculations, data processing, automated reasoning or decision making« algorithms in fact automate technology's computing (IIG 2023). To summarize, computerized algorithms are »structured decision-making processes that *automate* computational procedures to generate decisional outcomes on the basis of data inputs« (Gal 2018, 64; emphasis by author). They are the technical foundation for many modern technologies that we use every day. Digital platforms, such as Google, Netflix or Instagram, are just a few examples of this.

In this sense, digital, algorithm-based technologies, such as AI, could initially be ascribed a certain degree of agency as these systems can partly perform the task assigned to them independently or without further human input. However, a glance at the interrelation of agency and autonomy reveals that those processes which *prima facie* may appear to be a kind of autonomous action of the technology turn out to be merely a form of precise – albeit by now highly complex – automation. Within the philosophical discourse it is widely acknowledged that an entity's ›agency‹ manifests its capacity to act or exert power on its own (Schlosser 2019). Akin to the concept of ›autonomy‹, it denotes a comparative state of mind – i.e., a state that can be gradually increased or decreased. Depending on a person's or system's agency – meaning their ability »to act on the beliefs and values they hold« (Prunkl 2022, 3) – these agents can be considered more or less autonomous or self-determined. Hence, both philosophical concepts agency as well as autonomy are intertwined. Yet, their relation indicates that a sufficient degree of agency enables autonomy in the first place. Against this background, it can therefore be presumed that algorithm-based technologies – to a certain extent – operate as actors. Yet, this does not inevitably imply that they are autonomous, since the computational processes or actions of AI, for example, are based on previously defined decision paths and not on the rationality or free will of the technology in question. In this sense, the technology lacks its very own impetus and reasoned deliberation to perform an action in question. Hence, algorithm-based systems do not operate out of intrinsic or extrinsic motivations, but because they have been created by humans to process these tasks by means of specialized algorithms. Although algorithms can ultimately provide certain recommendations

– e.g., which movie might be of interest to us or whether a melanoma is good or malignant – such ›decisions‹ or ›actions‹ are, hence, not based on the free will or the rationally guided deliberation of the processing technical system. Rather, the machine merely executes a sequence of programmed, previously well-defined computational steps that have been inscribed into the technology by computer scientists. The »decisional parameters and rules for weighing them«, however, are originally »set by the algorithm's designer« (Gal 2018, 64f.). So, algorithm-based technology can never be fully autonomous. Or as Russell, a leading AI researcher and professor of computer science at the University of California (Berkeley), stated: »[…] machines, unlike humans, have no objectives of their own, we give them objectives to achieve. In other words, we build machines, feed objectives into them, and off they go. The more intelligent the machine, the more likely it is to complete that objective« (Russell 2019).

So concludingly, computer scientists create the functional spectrum of mechanical autonomization by programming tailored algorithms, which in turn have enormous effects or backlashes on us. Regarding the rapid developments in the field of algorithm-based technology, it is therefore imperative to further question how algorithm-driven automation affects us citizens in our democratic autonomy.

3. ›Algorithmization‹ of Political Communication: Implications for public opinion-forming and decision-making

In the following, these backlashes will be systematically assessed regarding the above-mentioned conditions of joint public reasoning according to Habermas. Central to this, however, are less the informatics underlying algorithms than their political effects and objectives: What functions, for example, do algorithms fulfill in the context of political communication? Do they serve as a decision guidance for human action, as they are often portrayed (e.g., Liang et al. 2022), or do they gradually disenfranchise humans in the context of their political decision-making?

The subtle ›algorithmization‹ of political communication can be witnessed in various settings: Political actors – such as elected officials or representatives of interest groups and NGOs – for example, use technological proxies in the form algorithmically (semi-)automated chat bots to disseminate their messages in the most cost- and time-efficient way (Woolley/Howard 2016). Moreover, algorithms are utilized to provide people with different messages.

By means of so-called micro-targeting strategies, people are given different information depending on their algorithmically calculated political interests, demographic and social profiles (Zarouali et al. 2022). Further, DL-algorithms are used to create or replace the image of one person in videos and other digital media, for example a high-ranking member of parliament or a military. So-called deep fakes are generated thanks to advanced algorithms. So, algorithms can be utilized for strategically controlling political communication and thereby political opinion making by the distortion of the supposed public opinion in various ways. Acting as a digital, stochastic calculating *spin doctor*, they deeply interfere in the process of shaping political opinion. Capable of shattering trust in politics and the overall facticity of information, they ought not to be underestimated (Whyte 2020). Algorithms do not only (pre-)select »what information is considered most relevant to us« (Gillespie 2014, 167) and then decide with whom which information shall be shared, they also generate and curate information by themselves. Mostly neither with our knowledge nor consent, algorithms manage our interactions on social networking platforms, our personal preferences, and whether we participate in a political discourse or not. In other words, algorithms do not just alter political communication but create the digital public sphere around us, in which we can exchange ideas with others. The amphitheater-like structure of the public sphere, in which politically relevant issues are debated and solutions sought together, however, is giving way to algorithmically computed filter bubbles and echo chambers – i.e., parallel ›public spheres‹ in which people no longer exchange views with dissidents but with like-minded people. This algorithm-driven fragmentation of the public sphere and the resulting impediment to public deliberation are only some – yet the most prominent – destructive consequences Habermas analyzes in his latest work *A New Structural Transformation of the Public Sphere and Deliberative Politics* (2022).

With this power of governing the flows of information, algorithms change political communication drastically. By shaping communication channels and their respective content, algorithms deeply intervene in the process of public opinion- and decision-making. In this sense, the algorithmization of political communication equates to a radical intrusion into democratic processes in general and citizens' ability to truly exercise their democratic autonomy in particular; as the state of information and the algorithmically induced fragmentation of the public sphere have a very significant influence on how and whether people can and want to engage in discourse: algorithms inevitably impact the *accessibility* and *inclusivity* of discourse, the *equality* of contributions and oppor-

tunities for participation, the *sincerity* of all participants, and the *informality* of exchange. Due to algorithms, the supposedly invisible becomes visible and the supposedly visible becomes invisible. What may feed new, rather minority voices and ideas into the discourse, may also cancel out opinions or deliberately bring them into focus.

To conduct a more nuanced analysis of the algorithm-driven expansion or restriction of democratic autonomy and action, it seems useful to first take a more nuanced look at algorithms' various functions in dealing with information. Drawing on Sauerwein et al. (2022) research, three key functions can be distinguished: the (I) curation and filtering of existing content, the (II) moderation of specific content, and finally the (III) verification of this content. In the following, the effects of all these categories on people's democratic autonomy or their jointly reasoned self-rule will be considered.

3.1 Curation and filtering of content

The curation and filtering of information – i.e., finding, compiling and organizing relevant content – is one of the most discussed functions of algorithms in the context of political communication (cf. Berman/Katona 2020). According to Berman and Katona (2020, 298f.), the literature on curation algorithms initially focused on the economic design and the impact of algorithmic curation and filtering on consumers. Other researchers examined the impact on user consumption behavior and any change in terms of the streamed information's quality (Athey et al. 2017). The citizen perspective, on the other hand, lately received attention with studies on trends of polarization of discourse triggered by curating algorithms, increasing authorship by users or more active participation in online discussions, and the spread of political ideologies through pre-filtering algorithms (Cho et al. 2020).

However, the observed and expected effects of this seem to be ambivalent. At first, the potential of algorithmic curation gives rise to hopes of a substantial expansion of democratic autonomy: By reliably processing the large amount of information and data available today according to appropriate criteria, information that is relevant for discourse or common reasoning can be identified, filtered, and prepared for further reflection. By means of their advanced and robust pattern recognition, algorithms could not only distinguish valid contributions from illegitimate ones – for example, hate speech and other defamation and thus, according to Habermas, illegitimate content in political discourse – but contributions could also be pre-structured

into specific argumentation patterns or core statements. In this way, factual arguments could be filtered out of vast piles of data and fed into the discourse in a way that is prepared for further discussion. In addition to possible strands of argumentation, it would also be possible to cluster interest groups and the premises or demands underlying the contributions, which could be hierarchized in a further step by means of algorithms according to their logic and stochastically mapped according to their agreement in the discourse. Such possibilities would not only make it easier for citizens to obtain relevant information more quickly, but also provide them with an already sensibly prefabricated basis for the further reflection process. Curating and filtering information through algorithms can not only simplify and optimize access to content relevant for legitimate, public opinion and decision-making, it also facilitates linking and confronting the participants and their ideas within discourse – regardless of their intellectual capabilities, individual language skills or eloquence etc. Advanced algorithms in natural language models, for example, are already capable of reading out written text, transforming spoken words into text and/or translating it into a desired language. Such accessibility, in turn, would allow more people – especially minorities, such as blind or illiterate people – to participate equally in political discourse and make their voices heard. Citizen's jointly reasoned self-rule and the public reasoning as such, it could be argued, would be simplified by the algorithm-based automation of these curations; the exercise of the democratic autonomy of citizens would thus be facilitated.

And yet, automating the curation of information by delegating it to algorithms does not necessarily lead to greater inclusion, participation, and a higher quality of political discourse. After all, the curation of information is always necessarily connected with the evaluation, selection or filtering of information, from which some pitfalls for democratic autonomy can be derived: Saurwein et al. (2022), for example, identify several risks associated with technically automated filtering and curation. For example, the emergence of errors and unwanted selections by algorithms. If, for example, admissible contributions were filtered in advance and excluded from the discourse based on incorrectly placed markers or keywords, this would not only violate the process condition of inclusion and accessibility, but also the participation rights of citizens. In this respect, their autonomy would be unjustifiably restricted by automation via algorithms. The targeted manipulation of citizens during the public opinion and decision-making process can also be seen as a significant restriction of democratic autonomy. This is because the parameters according

to which the responsible algorithms classify content as relevant or irrelevant are usually unknown to us. This intransparency or opacity of algorithm-based filtering and preselection of information ultimately leads to a downright distortion of public opinion formation – resulting in cutting citizens' self-determination. Citizens could no longer reason together, on an equal footing, when they have such a variety of information at their disposal. According to the current logic of algorithms, information flows are strongly tailored to individual needs and preferences. This logic is inevitably followed by information asymmetry, which makes reasoning on an equal footing, without constraints almost impossible. The emergence of such information and power asymmetries, on the other hand, can be additionally reinforced by strategies such as micro-targeting already mentioned.

By consciously and specifically controlling the flow of information through micro-targeting – i.e., addressing tailor-made messages, which is geared toward addressing pursuable or mobilizable citizens (Kruschinski/Haller 2017, 3) – information can be disseminated without passing through the public space and made accessible to all. This unequal or imbalanced exchange of information, in turn, can lead to a distortion of reality and thus of perceived public opinion. So political microtargeting by means of algorithms, however, does not necessarily lead exclusively to a reduction of democratic autonomy. Citizens can also be encouraged to participate in discourse when information is curated and filtered accurately and correctly. Tailored messages might be more relevant for users, helping them to better keep up with central political issues and arguments, which might amplify the effects of better voter turnouts and a more active citizenry exercising their democratic autonomy. In addition to this potential for mobilizing citizens, their political knowledge and general informedness could also be improved through algorithm-driven subject-oriented curation, which in turn could improve the fundamental quality of shared reasoning.

3.2 Moderation of specific content

As Gorwa et al. (2020, 1) state »[a]utomated hash-matching and predictive machine learning tools – what we define […] as algorithmic moderation systems – are increasingly being deployed to conduct content moderation at scale […] for user-generated content«. By means of these algorithmic moderation tools, problematic, toxic, or hostile statements such as hate speech can effectively be identified, deleted and prevented. With the increased use of algorithms for

content moderation, a growing body of research examines the political, social, and economic impact of this function of algorithms (cf. Kaye 2019).

Drawing on Gorwa et al. (2020, 3), we can distinguish between two moderating systems: Those, who decide about the visibility of accounts and their content (hard moderation) and systems utilizing rather soft moderation-techniques, such as recommendation systems or nudging by design. The political power of algorithms, however, becomes particularly evident in its function of hard content moderation: through the moderation of content – i.e., the technically automated decision-making as to which information is permitted in the discourse and which statements must be deleted or censored – a wide-reaching power but also responsibility is transferred from the people to the technology they design. Similar to curating, this second function of algorithms decides how and what information we perceive and what will be included into the political discourse. Against this background, algorithms also function as a new mechanism of gatekeeping. What traditionally has been performed by humans now is taken over by algorithms. Accordingly, algorithms are in charge of the content that is »visible and therefore noticeable to users, and influence the diversity of content that is consumed« (Stark et al. 2020, 10).

The use of these algorithm-based, automated techniques for hard and soft content moderation, on one hand, can potentially improve the level of discourse and in this way to support the legitimacy of political opinion and decision-making. Because algorithms can effectively contribute to supporting the processes of the jointly reasoned self-rule by allowing or excluding content according to the discourse-ethical conditions of these processes. Thereby, they create a democratic or ›safe space‹ in which everyone can freely express their respectful and legitimate opinion. In this way, it might be possible to mobilize citizens to join the discourse or public reasoning which feared the hostilities and hatred within unmoderated communication. In principle, they undermine the authorship of the people who have written hate comments and other illegitimate contributions, but in terms of democratic theory, those statements were not permitted in political discourse or were part of democratic autonomy anyways. On the other hand, gatekeeping by algorithms also restricts us in our autonomy, ability to act and freedom of decision, as they restrict the information and options through their moderation and gatekeeping and thus define the framework for our actions.

However, if the criteria according to which algorithms moderate content correspond to the commercial logic of the platform economy (*clickbaiting*) – instead of the above-mentioned discourse ethical premises – algorithms will not

remove fake news and hate speech etc., but rather spread it throughout society and political dispute. Thereby, incorrectly coded moderation algorithms might undermine the »epistemic potential« of political discourse.

3.3 Verification of content

In the context of political communication, algorithms are also used to verify content. For this purpose, algorithms are developed for reviewing claims »with reference information in the form of facts in a knowledge base« (Huynh/Papotti 2019, 689). Hence, the ability, applicability, and performance of fact-checking algorithms of course is limited to the set of knowledge with which humans trained them and the pool of information which is available to them – without any human supervision or reinforcement.

In their function of verifying the objective accuracy, truth or factuality of information, algorithms again fulfill the role of a modern, digital gatekeeper that seems to be successively replacing the former control mechanisms, such as journalists and mass media. While algorithm-based platforms and search engines, such as Google or Bing, facilitate access to content and therefore expand human agency, algorithms specially trained to identify fake news, deep fakes and other false statements additionally ensure human access to verified information and knowledge. This, in turn, not only enables free accessibility to objective facts or information, but it also increases people's informed self-determination.

How dangerous disinformation and the targeted dissemination of false statements can be for the integrity and legitimacy of political processes has become something of a commonplace in discourse (Morgan 2018). Yet, disinformation campaigns not only polarize and radicalize political discourse, they affect personal autonomy *per se*. As Witzleb and Paterson (2020, 227) note, mis- as well as disinformation in its various forms is »harmful as it has the potential to disrupt our individual capacity for self-authorship and, as a consequence, our communal capacity for self-government«. False information can interfere with democratic autonomy, as both the sincerity and the honest solution orientation of the agents involved in political debates can no longer be ensured.

The implementation of algorithms to contain the spread of misinformation and disinformation and its threat for democratic society and citizens' democratic self-determination can therefore rather be seen as an autonomy supporting than a diminishing function.

4. Democratic Autonomy vs. Algorithms: Modelling the limitation and expansion of democratic autonomy

Based on the previous discussion of the algorithmization of political communication, finally, an attempt will be made to model the potential limitation and expansion of democratic autonomy by algorithms. To this end, the developed theses on the political implications of the (I) curation and filtering, the (II) moderation and (III) verification of content by algorithms will be systematized and summarized.

Considering the process conditions of joint deliberation and democratic autonomy, the elaborated merits and drawbacks can be presented as follows. The table is divided into the preconditions of democratic autonomy and the discourse conditions of joint reasoning, on one hand, as well as the functions of algorithms in the context of political communication, on the other. The markings indicate which conditions are endangered by which function:

Table 1: Synthesis of limits and opportunities for democratic autonomy due to algorithms

Condition **Function**	Accessibility & Inclusivity	Equality of Contributions & Participation	Sincerity of all Participants	Informality of Exchange
Curation & Filtering				X
Verification				
Moderation		X		X

In conclusion, it became evident that the individual functions of algorithms point to their power as new control mechanisms of the information flow and gatekeepers of discourse. The algorithmically controlled automation of information and communication, it was shown, creates the danger of »*shortcuts*«. The arduous process of jointly public reasoning – i.e., the active exercise of democratic autonomy, which is tied to rather demanding conditions – could possibly be circumvented or, in some aspects, even replaced by algorithms. Like many democratic theorists, Lafont – a former student of Habermas – warns of such cutbacks, as they have occurred throughout history (Lafont 2020). The participation of citizens and the free exercise of their self-determination should under no circumstances be undermined by supposedly efficient shortcuts. The »ideal of self-government« as vital, »participatory aspect of democracy« (Lafont 2020, 7), threatens to be undermined by taking supposedly easier or more efficient paths.

By contrast, the technical autonomization of fundamental, democratic processes, such as one's own walk to the ballot box or the struggle to find solutions to political issues, can be seen as such a shortcut. The algorithm-based processing of information necessary for reflective participation in public decision-making and opinion-forming could already be interpreted as the preliminary stage of such a development. Political alienation and the impediment of true democratic autonomy are the consequences of such shifts. Or as Lafont claims: »Democratic *participation* in decision-making is essential to prevent an *alienating disconnect* between the political decision to which citizens are subject and their political opinions and will« (Lafont 2020, 22f.). Algorithms, it had been shown, appear in both forms: On the one hand, they serve us as decision-making aids by facilitating access to objective facts and consensus-oriented input. They promote the inclusion of the other and in particular of those who think differently. Whereas, on the other hand, algorithms also force the exclusion of the other and of those who think differently by creating deep fakes or by moderating content in discourse in a click-oriented rather than a consensus-oriented manner. In this way, they no longer function as a decision-making aid in political interaction, but as a distorting mirror of public opinion and perception, which makes discussion at eye level and the free exercise of democratic self-determination by all almost impossible. The danger of manipulation and thus, to a certain extent, the ›disenfranchisement‹ of citizens grows with every use of algorithms and bots to deliberately change the climate of opinion. Citizens, with their individual opinions, rights and

political views, are increasingly relegated to the background. They are reduced to a number to be recorded statistically.

Algorithms, on the other hand, are to be commended if they improve the quality of public reasoning. For example, through the targeted verification of content, the identification and subsequent filtering of toxic statements and the clustering of certain strands of argument. Algorithms should therefore be used to improve the level of political debate and thus also protect the ideal of democratic self-determination. On the other hand, they should not be used to subtly and implicitly as well as explicitly force citizens to forego their right of political self-determination or even to no longer be able to exercise it. Rather, they should be used »*to force the political system to take ›the long road‹ of properly involving citizenry*« (Lafont 2020, 159).

5. Concluding Remarks

Thanks to the ever-growing availability of data and the increasing power of algorithms, political practices and processes are changing. But what are the implications of the increasing use of algorithms for people's political or democratic autonomy? Do algorithms have an empowering or disempowering effect on us citizens?

The paper aimed to address this question by drawing on references from the philosophy of technology as well as from the theory of democracy. In course of this, the philosophical concept of autonomy was extended to the political idea of so-called democratic autonomy, as elaborated by Habermas. The paper, further, offered a short note on the importance of algorithms as key mechanisms for technical autonomation. It became clear that algorithms have a profound impact on the way people discuss current affairs and engage with each other as well as politics.

In conclusion, the paper argues that algorithms can be a threat to citizens' autonomy and their right to informational self-determination at the individual level, but they can also be challenging to the discursive process of opinion- and will-forming at the structural level. In both cases, however, algorithms decisively determine the basis on which humans' political decisions are made or an action is taken. Regarding the systematization of algorithms' functions and their impact on the preconditions of jointly public reasoning, the paper found that the (3.1) curation and filtering by algorithms infringes the necessary informality of exchange. Whereas algorithms (3.2) verifying content are empower-

ing humans in their ability to participate in discourse or the joint public reasoning, as it contributes to the improvement of the discourse's quality and thus counteracts the chilling effects and the radicalization and polarization of the discourse. The (3.2) content moderation by algorithms, in contrast, might diminish both the informality of exchange as well as the equality of contributions and participation in discourse.

As the paper noted, human autonomy has become a theme across various guidelines and principles on the responsible use and design of algorithms. However, according to Sætra, Borgebund and Coeckelbergh (2022, 804), there is a tendency within AI research to adopt the concept of democracy and its associated norms merely as buzzwords – without referring to their further political and historical significance. The paper's objective has been to shed light on the political dimension of the various effects the growing power of algorithms hold for humans' political or democratic autonomy. Against this background, fundamental questions remain untouched, such as whether political discourse will be obsolete one day because the better argument or the better solution for political conflicts could be found algorithmically. Are algorithm-based automations, such as finding political solutions through mathematical calculations instead of the deliberative ideal of communicative negotiation, be permissible in terms of democratic theory or should they be rejected as destructive shortcuts? These and many other questions remain to be researched in the future.

References

Anderson, Joel. 2011. »Autonomy, Agency, and the Self.« In *Jürgen Habermas. Key Concepts*, edited by Barbara Fultner, 91–114. Durham: Acumen.

Athey, Susan, Markus M. Mobius, Jeno Pal. 2017. »The impact of aggregators on internet news consumption.« Working paper, Stanford University, Stanford, CA.

Batarseh, Feras A., Ruixin Yang. 2020. *Data Democracy: At the Nexus of Artificial Intelligence, Software Development, and Knowledge Engineering*. Academic Press.

Belford, Geneva G., Allen Tucker. 2023. »Computer Science.« *Encyclopedia Britannica*. Accessed April 23, 2023. https://www.britannica.com/science/computer-science/Algorithms-and-complexity.

Berman, Ron, Zsolt Katona. 2020. »Curation algorithms and filter bubbles in social networks.« *Marketing Science*, 39(2): 296–316.

Chiodo, Simona. 2022. »Human autonomy, technological automation (and reverse).« *AI & society* 37: 39–48.

Cho, Jaeho, Saifuddin Ahmed, Martin Hilbert, Billy Liu, Jonathan Luu. 2020. »Do search algorithms endanger democracy? An experimental investigation of algorithm effects on political polarization.« *Journal of Broadcasting & Electronic Media* 64 (2): 150–172.

Eisenegger, Mark. 2021. »Dritter, digitaler Strukturwandel der Öffentlichkeit als Folge der Plattformisierung.« In *Digitaler Strukturwandel der Öffentlichkeit: Historische Verortung, Modelle und Konsequenzen*, edited by Mark Eisenegger, Marlis Prinzing, Patrik Ettinger, Roger Blum. Springer VS. 17–39.

Gal, Michael S. 2018. »Algorithmic challenges to autonomous choice.« *Michigan Technology Law Review* 25: 59–104.

Gillespie, Tarleton. 2014. »The relevance of algorithms.« *Media technologies: Essays on communication, materiality, and society* 167 (2014): 167–193.

Goldschmitt, K. E., Nick Seaver. 2019. »Shaping the Stream: Techniques and Troubles of Algorithmic Recommendation.« In *The Cambridge Companion to Music in Digital Culture*, edited by Nicholas Cook, Monique M. Ingalls, David Trippett. Cambridge University Press, 63–81. DOI: https://doi.org/10.1017/9781316676639.006.

Gorwa, Robert, Reuben Binns, Christian Katzenbach. 2020. »Algorithmic content moderation: Technical and political challenges in the automation of platform governance.« *Big Data & Society* 7 (1). DOI: https://doi.org/10.1177/2053951719897945.

Habermas, Jürgen. 2009. *Zwischen Naturalismus und Religion. Philosophische Aufsätze*. Frankfurt a.M.: Suhrkamp.

Habermas, Jürgen. 2019 [1992]. *Faktizität und Geltung*. Frankfurt a.M.: Suhrkamp.

Habermas, Jürgen. 2022. *Ein neuer Strukturwandel der Öffentlichkeit und die deliberative Politik*. Berlin: Suhrkamp.

Heßler, Martina. 2019. »Technik und Autonomie.« In *Autonome Systeme und Arbeit. Perspektiven, Herausforderungen und Grenzen der Künstlichen Intelligenz in der Arbeitswelt*, edited by Hartmut Hirch-Kreinsen, Anemari Karačic. Bielefeld: Transcript, 247–274.

Huynh, Viet-Phi, Paolo Papotti. 2019. »A benchmark for fact checking algorithms built on knowledge bases.« *Proceedings of the 28th ACM Interna-*

tional Conference on Information and Knowledge Management, November 2019, 689–698.

International Institute in Geneva. 2023. *Understanding Algorithms in Computer Science*. Accessed April 23, 2023. https://www.iig.ch/en-en/blog/computer-science/algorithm-computer-science-definition-and-understanding.

Kaye, David. 2019. *Speech Police: The Global Struggle to Govern the Internet*. New York, NY: Columbia Global Reports.

Kruschinski, Simon, André Haller. 2017. »Restrictions on data-driven political micro-targeting in Germany.« *Internet Policy Review* 6 (4): 1–23. DOI: https://doi.org/10.14763/2017.4.780.

Kneuer, Marianne. 2019. »E-Democracy.« In *Handbuch Digitalisierung in Staat und Verwaltung*, edited by Tanja Klenk, Frank Nullmeier, Göttrik Wewer. Springer Fachmedien: Wiesbaden, 267–277.

Lafont, Cristina. 2020. *Democracy without shortcuts: A participatory conception of deliberative democracy*. Oxford University Press.

Laitinen, Arto, Otto Sahlgren. 2021. »AI systems and respect for human autonomy.« *Frontiers in artificial intelligence* 4: 1–14.

Liang, Garston, Jennifer F. Sloane, Christopher Donkin, Ben R. Newell. 2022. »Adapting to the algorithm: how accuracy comparisons promote the use of a decision aid.« *Cognitive research: principles and implications* 7 (1):14. DOI: https://doi.org/10.1186/s41235-022-00364-y.

Lovett, Adam, Jake Zuehl. 2022. »The Possibility of Democratic Autonomy.« *Philosophy & Public Affairs* 50 (4): 467–498.

Morgan, Susan. 2018. »Fake news, disinformation, manipulation and online tactics to undermine democracy.« *Journal of Cyber Policy* 3 (1): 39–43.

Paulin, Alois. 2020. »An overview of ten years of liquid democracy research.« *The 21st Annual International Conference on Digital Government Research* June 2020: 116–121.

Pentenrieder, Annelie. 2021. »Algorithmen erklärt euch.« In *In digitaler Gesellschaft*, edited by Kathrin Braun, Cordula Kropp. transcript Verlag, 53–69.

Pohlmann, Rosemarie. 2010: »Autonomie«. In *Historisches Wörterbuch der Philosophie*, edited by Joachim Ritter, Karlfried Gründer, Gottfried Gabriel. Berlin: Schwab.

Prunkl, Carina. 2022. »Human autonomy in the age of artificial intelligence.« *Nature Machine Intelligence* 4 (2): 99–101.

Quante, Michael. 2013. »Autonomie«. In *Lexikon Philosophie. Hundert Grundbegriffe*, edited by Stefan Jordan, Christina Nimtz. Stuttgart: Reclam, 47–48.

Rawls, John. 1996. *Political Liberalism*. New York: Columbia University Press.

Richardson, Henry S. 2003. *Democratic Autonomy: Public Reasoning about the Ends of Policy*. Oxford: University Press.

Russell, Stuart. 2019. »How to stop Superhuman A.I. before it stops us.« *New York Times*. Accessed April 24, 2023. https://www.nytimes.com/2019/10/08/opinion/artificial-intelligence.html.

Sætra, Henrik Skaug. 2021. »A typology of AI applications in politics.« In *Artificial Intelligence and Its Contexts: Security, Business and Governance*, edited by Anna Visvizi, Marek Bodziany. Springer VS, 27–43.

Sætra, Henrik Skaug, Harald Borgebund, Mark Coeckelbergh. 2022. »Avoid diluting democracy by algorithms.« *Nature Machine Intelligence* 4 (10): 804–806.

Saurwein, Florian, Charlotte Spencer-Smith, Jaro Krieger-Lamina, et al. 2022. »Social-Media-Algorithmen als Gefahr für Öffentlichkeit und Demokratie: Anwendungen, Risikoassemblagen und Verantwortungszuschreibungen.« In *Digitalisierung und die Zukunft der Demokratie. Beiträge aus der Technikfolgenabschätzung*, edited by Alexander Bogner, Michael Decker, Michael Nentwich, Constanze Scherz. Baden-Baden: Nomos, 243–256.

Sankaran, Supraja, Chao Zhang, Mathias Funk, et al. 2020. »Do I have a say? Using conversational agents to re-imagine human-machine autonomy.« *Proceedings of the 2nd Conference on Conversational User Interfaces*, 1–3.

Schlosser, Markus. 2019. »Agency.« In *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2019/entries/agency/.

Stark, Birgit, Daniel Stegmann, Melanie Magin, Pascal Jürgens. 2020. »Are algorithms a threat to democracy? The rise of intermediaries: A challenge for public discourse.« *Governing Platforms. Algorithm Watch*. Accessed April 23, 2023. https://algorithmwatch.org/en/wp-content/uploads/2020/05/Governing-Platforms-communications-study-Stark-May-2020-AlgorithmWatch. pdf.

Thimm, Caja, Thomas Christian Bächle. 2019. »Autonomie der Technologie und autonome Systeme als ethische Herausforderung«. In *Maschinenethik: Normative Grenzen autonomer Systeme*, edited by Matthias Rath, Friedrich Krotz, Matthias Karmasin. Springer VS, 73–87.

Verständig, Dan, Christina Kast, Janne Stricker, Andreas Nürnberger (eds.). 2022. *Algorithmen und Autonomie: Interdisziplinäre Perspektiven auf das Verhältnis von Selbstbestimmung und Datenpraktiken*. Opladen: Barbara Budrich.

Weber, Max. 2006 [1948]. *Wirtschaft und Gesellschaft*. Paderborn: Voltmedia.

Weithman, Paul. 2011. »Convergence and political autonomy.« *Public Affairs Quarterly* 25 (4): 327–348.

Whyte, Christopher. 2020. »Deepfake news: AI-enabled disinformation as a multi-level public policy challenge.« *Journal of cyber policy* 5 (2): 199–217.

Witzleb, Norman, Moira Paterson. 2020. »Micro-targeting in Political Campaigns: Political Promise and Democratic Risk.« In *Data-Driven Personalisation in Markets, Politics and Law*, edited by Uta Kohl, Jacob Eisler. Cambridge, 223–240.

Woolley, Samuel C., Philip N. Howard. 2016. »Political communication, computational propaganda, and autonomous agents: Introduction.« *International journal of Communication* 10: 4882–4890.

Zarouali, Brahim, Tom Dobber, Guy De Pauw, Claes de Vreese. 2022. »Using a personality-profiling algorithm to investigate political microtargeting: assessing the persuasion effects of personality-tailored ads on social media.« *Communication Research* 49 (8): 1066–1091.