



Political machines: a framework for studying politics in social machines

Orestis Papakyriakopoulos¹

Received: 28 April 2020 / Accepted: 9 March 2021
© The Author(s) 2021

Abstract

In the age of ubiquitous computing and artificially intelligent applications, social machines serves as a powerful framework for understanding and interpreting interactions in socio-algorithmic ecosystems. Although researchers have largely used it to analyze the interactions of individuals and algorithms, limited attempts have been made to investigate the politics in social machines. In this study, I claim that social machines are per se political machines, and introduce a five-point framework for classifying influence processes in socio-algorithmic ecosystems. By drawing from scholars from political theory, I use a notion of influence that functions as a meta-concept for connecting and comparing different conceptions of politics. In this way, I can associate multiple political aspects of social machines from a cybernetic perspective. I show that the framework efficiently categorizes dimensions of influence that shape interactions between individuals and algorithms. These categories are symbolic influence, political conduct, algorithmic influence, design, and regulatory influence. Using case studies, I describe how they interact with each other on online social networks and in algorithmic decision-making systems and illustrate how the framework is able to guide scientists in further research.

Keywords Political machines · Social machines · Sociotechnical systems · Algorithmic bias · Tech policy · Cybernetics · Artificial intelligence · Design · Algorithms

1 Introduction

The consolidation of the internet as society's main communication network, the development of hardware with increasing computational efficiency, the invention of the smartphone, and general social datafication (Mayer-Schönberger and Cukier 2013) have created new communication and information processes that impact all aspects of society. Algorithmic decision-making systems, artificially intelligent algorithms, online social networking platforms, data sharing systems and predictive tools invade everyday life, leading to the transformation of human behavior, socialization, economic markets and political conduct.

One of the most intensive reformations caused by technological advancement is the degree of digitization of social processes. The new computational and storing capabilities lead to the constant transformation of individual behavior

into metadata being processed and stored, while integrated internet access on any type of device facilitates permanent information exchange. Hence, algorithms, society and communication are coupled in complex and continuous ways, generating new forms of socio-algorithmic ecosystems. The founder of the world wide web, Tim Berners-Lee, defined such ecosystems, in which individuals and algorithms participate and interact, as social machines (Berners-Lee and Fischetti 2001; Shadbolt et al. 2019). Social machines are not per se machines, nor do they depict mechanistic deterministic phenomena. On the contrary, they are systems of systems (Hendler and Mulvehill 2016), in which humans and algorithms are detached from their materiality, forming complex interaction patterns. Online social networks, algorithmic decision making (ADM) systems, search engines, are all types of social machines, with individuals, software and hardware constantly interacting and resulting in emergent system states.

✉ Orestis Papakyriakopoulos
orestis@princeton.edu

¹ Center for Information Technology Policy, Princeton University, Princeton, USA

1.1 Motivation

Although researchers have largely analyzed socio-algorithmic systems as social machines (Smart and Shadbolt 2015; Shadbolt et al. 2019; Hendler and Berners-Lee 2010; Buregio et al. 2013; Cristianini and Scantamburlo 2019), no efforts have been made to analyze the politics of these systems under a similar framework. Studies predominately focus on investigating the politics of separate parts of these systems, such as legislation, algorithmic function, or user behaviour. Therefore, a significant knowledge gap exists in understanding how political processes actually interact and impact each other, as well as how they transform the systems from a holistic perspective. For example, there are open questions about how changes in recommendation systems on social media platforms such as Facebook and YouTube influence the circulated political content, or how web mapping services such as Google Maps alter traffic patterns and user consuming habits.

The understanding of political aspects in cases such as the above is necessary for three reasons: (1) There is a need for additional and holistic scientific knowledge on how algorithms influence society since most of the studies are case specific and face serious limitations because of data bias and systems' opacity. (2) Legislation seeks ways to intervene in existing socio-algorithmic systems, to mitigate unjust, unethical, and illegal outcomes, but there is no framework to guide such interventions. (3) System designers are interested in understanding how the insertion of a technical component, or the existence of a regulation might influence the function of their socio-algorithmic ecosystem directly or indirectly. This study develops the theoretic foundation for answering questions such as the above, by introducing a framework that classifies the politics of socio-algorithmic ecosystems. The framework reduces the existing complexity of the interpretation of political processes in social machines. It achieves that by seeking the answer to the following research questions:

1.1.1 RQ1: How can researchers analyze and classify political processes in social machines from a systemic perspective?

1.1.2 RQ2: How can researchers use the above framework as a guide for understanding socio-algorithmic ecosystems?

1.2 Original contributions

- I study social machines under a cybernetic framework, unraveling the properties of the systems. I show that

social machines can be a valuable ecosystem-agnostic tool, supporting researchers when applying and testing their scientific narratives.

- I introduce the framework of political machines when investigating politics in social machines. The framework promotes the normative statement that technology is not neutral participant in society. In turn, algorithmic implementations radically transform sociopolitical function in unexpected ways.
- I describe the political machines framework, which categorizes political processes in five main categories of influence. The framework adopts a notion of influence that functions as a meta-concept for connecting and comparing different conceptions of politics. By the use of cybernetics, it is able to advance understanding of complex political phenomena in socio-algorithmic ecosystems.
- By presenting two case studies, I illustrate how the developed framework can guide scientists in further research.

2 Background and related work

2.1 Social machines

Social machines are a paradigm for investigating, evaluating, and understanding socio-algorithmic ecosystems, largely influenced by computer scientific thought. It emerged as a scientific solution for dealing with the excessive social datafication and increasing interconnectedness of social and technological processes (Hendler and Berners-Lee 2010). As a scientific model, it aims to unify computational, technological, and social processes under the same framework (Buregio et al. 2013), supporting explanations that transcend the traditional boundaries set by scientific disciplines. In social machines, both individuals and technology are participants of systemic processes (Smart and Shadbolt 2015). By adopting systems' theory, scientists are able to trace inputs, outputs, interactions, constraints, and states that shape a specific social machine (Meira et al. 2011), abstracting what is human and what is artificial. This, in turn, reduces the complexity when studying phenomena and facilitates the practical understanding of socio-algorithmic ecosystems.

2.2 Social machines and other approaches

In this sense, social machines' contribution is similar to other approaches such as the actor-network theory (ANT) (Latour 2005), which aims to describe complex sociotechnological processes by placing humans and technology on the same level. Such frameworks aim to assist researchers in exploring such ecosystems, rather than to provide structured theoretical knowledge (Mol 2010). Nevertheless, the

framework of social machines possesses the following distinguishing features (Shadbolt et al. 2019): (1) It assumes that the system amplifies inputs, because of the pervasiveness and effectiveness of technology. (2) The studied ecosystem is a result of a design process, in contrary to, for example, networks in ANT, whose formation background is not investigated (Latour 2005). (3) The ecosystem has specific goals and features, which are emergent from the interactions of individuals and technology.

The above systemic perspective and the role of humans and technology are what distinguishes social machines from critical data studies and sociotechnical systems approaches. Both critical data studies and sociotechnical systems theories investigate the necessary epistemological concepts and questions that should be answered towards understanding and shaping the ethics, manifestation and influence of technology in the society (Iliadis and Russo 2016; Dalton, Taylor, and Thatcher 2016; Selbst et al. 2019; Norman and Stappers 2015). In contrast, social machines does not theorize specific cases but aims to assist such approaches in their scope by providing an ecosystem-agnostic framework that locates key participants in the interaction of technology and society and their interrelations from a descriptive perspective. This framework can be useful for scientists (from social science to engineering) for developing and evaluating narratives and conceptualizations of sociotechnical phenomena.

2.3 Social machines applications

Until today, multiple phenomena have been studied through the lens of social machines. Crowdsourcing platforms, online social networks, smart cities, internet of things applications and web-based communities are only some of the cases analyzed by the paradigm, mostly for computer-scientific goals (Shadbolt et al. 2019; Ahlers et al. 2016; Buregio et al. 2013; Martin and Pease 2013). Nevertheless, the definition and structured analysis of social machines is not a trivial task (Smart et al. 2014). Because of the various roles, objectives, and behaviours humans and technology adopt, it is difficult for researchers to frame absolute classification schemes of ecosystems that are quite distinct from each other. Towards that end, researchers have proposed different frameworks and taxonomies for evaluating social machines. For example, Buregio et al. (2013) classify social machines according to the contribution of the systems, their motivation, as well as who participates and how. Similarly, De Roure et al. (2015, n.d.) describe methodologies on what to observe in social machines and how, and Smart et al. (2014) investigate systems' prominent similarities and differences.

Despite the existing ambiguity, the framework offers new opportunities when dealing with socio-algorithmic ecosystems. Its ability to order scientific knowledge about complex systems has been deployed for understanding and

normatively evaluating how social machines influence society. Researchers have created theoretical foundations for evaluating social machines' contribution to society (Palermos 2017), as well as design principles for social machines from grass-roots and participatory movements perspectives (Papapanagiotou et al. 2018; Murray-Rust et al. 2018) (e.g. Fig. 1). Nevertheless, the vast amount of political processes in social machines have not been extensively studied under the paradigm, a gap that this study wants to bridge.

2.4 Investigating politics in social machines

Researchers have extensively analyzed various social machines, uncovering political properties and behaviours that constitute these systems. Nevertheless, the majority of the studies are case specific. A set of investigations analyzes the behaviour of social groups under the algorithmic influence, to uncover the algorithmic impact on public opinion and behaviour (Pariser 2011; Bakshy et al. 2015; Barberá et al. 2015). Other studies investigate information diffusion and opinion formation (Yang 2016; Faris et al. 2017; Stieglitz and Dang-Xuan 2013; Tufekci and Wilson 2012; Shahrezaye et al. 2019), while others dedicate themselves to understanding properties of spreading misinformation by real or artificial users (Del Vicario et al. 2016; Vosoughi et al. 2018; Ferrara et al. 2016; Papakyriakopoulos et al. 2020). In social machines such as social media or search engines, political actors also explicitly use platform tools for political campaigning. Given this, many researchers investigate how politicians place personalized advertisement on them and whether they influence the electorate and how (Endres 2016; Kruijkemeier et al. 2016; Schipper and Woo 2018).

The rise of social computation has also resulted in the generation of additional data sources that decision makers can exploit. Next to classical ADM systems used in areas such as mechanical and electrical engineering, and weather forecasting, new systems are being developed for various purposes such as autonomous vehicles, healthcare, economics, finance, employment, policing, and public administration, with human computation being of increasing importance (Dressel and Farid 2018; Barocas, Hardt, and Narayanan 2017). These systems exploit data-intensive algorithms and generate inferences about individuals that were not possible before. Most researchers developing these models are primarily interested in testing their efficiency and accuracy in comparison to other models and the human factor (Dressel and Farid 2018; Larson et al. 2018; Erickson et al. 2017). Other researchers focus on the ethical consequences of these methods: whether they are fair or discriminatory, how biases could be mitigated and how these systems should be regulated (Buolamwini and Gebru 2018; Dressel and Farid 2018; Bolukbasi et al. 2016; Barocas et al.

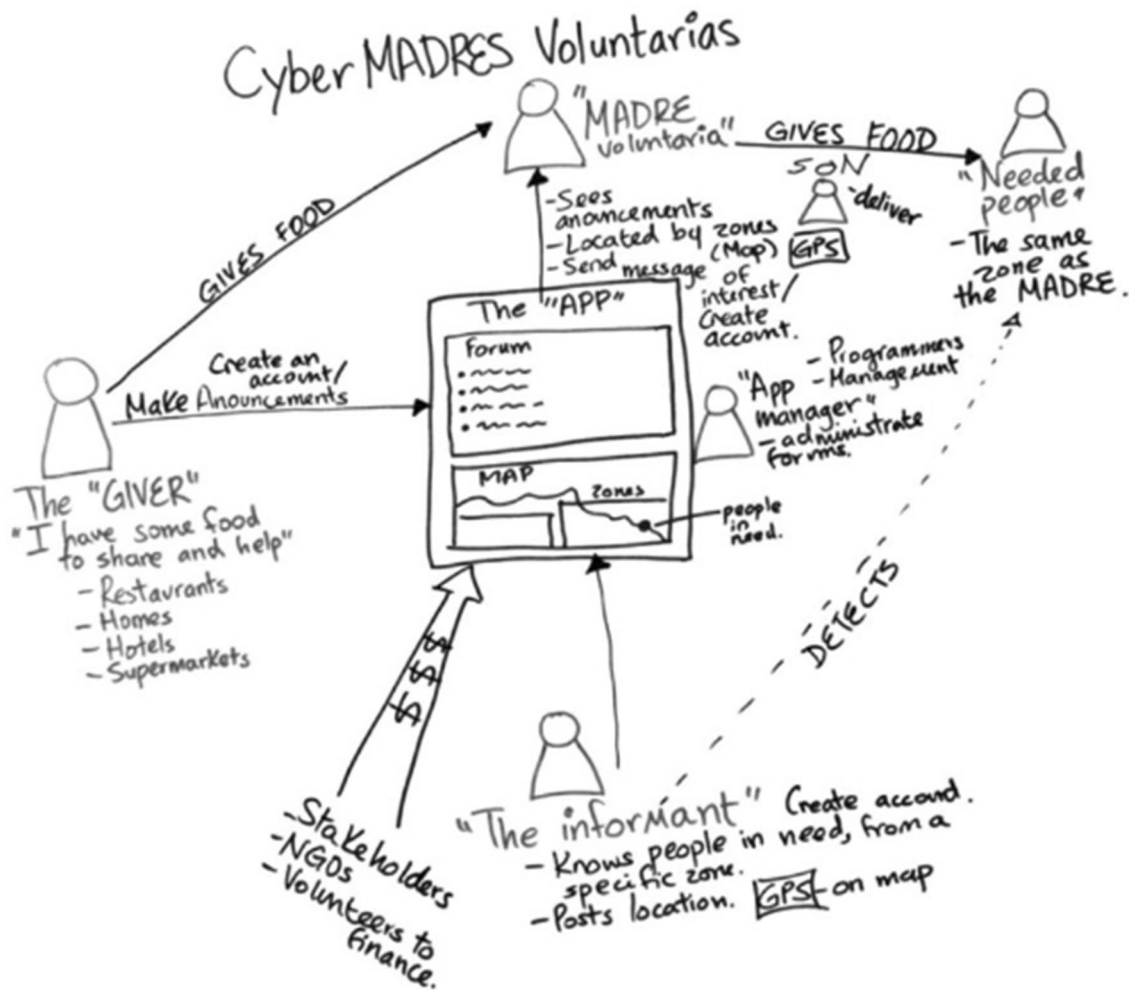


Fig. 1 The Cybermadres social machine sketch from (Murray-Rust et al. 2018). This Social Machine was designed for the activities of volunteers in Mexico, who collect excess food from restaurants and distribute it to people in need

2017; Kusner et al. 2017). Explaining and understanding algorithms is not only related to fairness but also to accountability and transparency. Given that legal frameworks, algorithmic design, and algorithmic influence interact with each other, scientists are trying to pose the correct questions to be answered. Towards that end, researchers analyze the interaction between data regulations, accountability, fairness, and the right to explanation (Wachter, Mittelstadt, and Russell 2017). Furthermore, they seek to uncover further cases of algorithmic bias (Mehrabi et al. 2019), and to detect further challenges in algorithmic fairness to form regulations and systems that conform to social imperatives (Chouldechova and Roth 2018; Bird et al. 2019).

Although researchers have analyzed many properties of politics in social machines, it is clear from the above that political processes are not only of high complexity but

appear in multiple parts of different social machines. This denotes that a terrain full of unknown unknowns exists in social machines, which wait to be discovered and understood. The most integrative approach towards this direction comes from the newly emerged field of machine behaviour (Rahwan et al. 2019), which focuses on the study of intelligent machines as a class of actors with particular behavioural patterns and ecology. In this way, the field tries to answer how the introduction of AI algorithms impacts society, as well as which sociopolitical factors shape the integration of algorithms in society. The framework of political machines that will be introduced later deals with similar questions but does not put algorithms in the epicenter. In contrast, it investigates a series of human, social and technological factors that constitute political processes in socioalgorithmic ecosystems.

3 Social machine cybernetics

Uncovering political processes in social machines requires a framework that provides a holistic overview of how socio-algorithmic ecosystems behave. The most prominent scientific theory that deals with systems and their behaviour is cybernetics (Wiener 2019). Cybernetics does not investigate systems just by looking at them as a set of inputs, outputs and interacting components. In contrast to other theories, it seeks to understand systems as they exist in a given environment, how their state changes according to the environment, what the systems' identity is, which constraints exist, what the processes of feedback, communication and control are that result in the transformation and self-organization of the system (Wiener 2019; Ashby 1957; Von Foerster 2007; Mead 1968; Rosenblueth, Wiener, and Bigelow 1943). Cybernetics does not reduce systems to things, but to ways of behaving. It does not question what is this thing? but what does it do? (Ashby 1957).

In cybernetics, communication is not reduced to human or animal communication, which corresponds to an explicit exchange of symbols and signs (Bateson 2006). Each type of interaction or impact between elements, systems, or the environment, can be treated as information that updates the related entities about the differences taking place (Novikov 2015; Ruesch et al. 2017), resulting in a higher order form of communication. The realization of difference, or else, is critical in cybernetics because it is able to uncover operators and operants in the system, i.e. what changes what and how? (Ashby 1957) By studying feedback loops, the cybernetician is able to uncover the purpose of elements and systems, as well as their specific structure and intrinsic organization (Rosenblueth et al. 1943).

3.1 Cybernetic applications and critique

Cybernetics as framework has already been applied for the investigation of the study of politics (Deutsch 1963) and the study of sociotechnological processes (Luhmann 1999; Lepskiy 2018). Karl Deutsch argued that cybernetics provides the necessary vocabulary for understanding political systems and power relations while being economic and empirically valid (Deutsch 1963). This is the case because politics can be seen as coordinating processes between system components. Similarly, Luhmann also argued for the coordinating role of technology in society, as well as stated that the function of media and technologies such as AI can be framed in terms of communication vocabularies, albeit of different nature than that of culture, leading to the creation of systems of heterogeneous components (Luhmann 1999). Despite the successful application of cybernetics for the study of sociotechnological ecosystems, there have also been strong

criticisms. One of the most prominent ones came from philosopher Jonas (1953), who argued that cybernetics assigns intentionality and goals in objects such as technological artifacts or systems that do not necessarily have one. Similarly, cybernetics reduces social processes to mechanistic descriptions. According to Jonas, these transformations are not justified, and lead to empty descriptions of the systems that only replicate the purposefulness and instrumentality that the researcher assigns to them.

I recognize the validity of Jonas' claims but argue that these features are not necessarily problems. First, the detachment of social and technological processes from their materiality leads to complexity reduction, which facilitates the easier understanding of studied phenomena. As I argued in the background of the study, the proposed framework of political machines does not aim to provide complete scientific knowledge. In contrast, it is a tool for assisting researchers and theorists when developing or testing their scientific narratives. Therefore, the use of the framework requires it to be complemented with further scientific theories to lead to a coherent and useful production of knowledge. Second, the ability to interpret the systems based on researcher's purposefulness underpins the fact that every knowledge is situated (Haraway 1988), as well as allows for system designers to intervene and reform systems according to their ends (Krippendorff 2019). This conforms with one of the objectives of the political machines framework, which is to be useful in normatively understanding, intervening and shaping socioalgorithmic ecosystems.

3.2 Social machines from a cybernetic perspective

In social machines, computability is not an exclusive right of the machines, nor is sociability an exclusive right of the humans. The cybernetic framework allows treating human behaviour as also computable, and the technological participation as sociable too. For example, human behaviour is projected into metadata fed into recommendation algorithms, deep learning models, or computer vision software. Similarly, the decision of an ADM System to hire or fire an individual replaces the human resources manager in a company's social network. Given that all these interactions are projected into forms of communication and control, they appear within the same cybernetic domain regardless of their initial materiality. What complements their regulation are the design frameworks that guide the behaviour of individuals and the application of technologies (Fig. 2). The design framework includes the values, infrastructure, exact algorithms, interfaces, regulations, and any other material or non-material property that constitutes the systems' elements behavior (Smart and Shadbolt 2015). For example, social media platform design is usually based on companies'

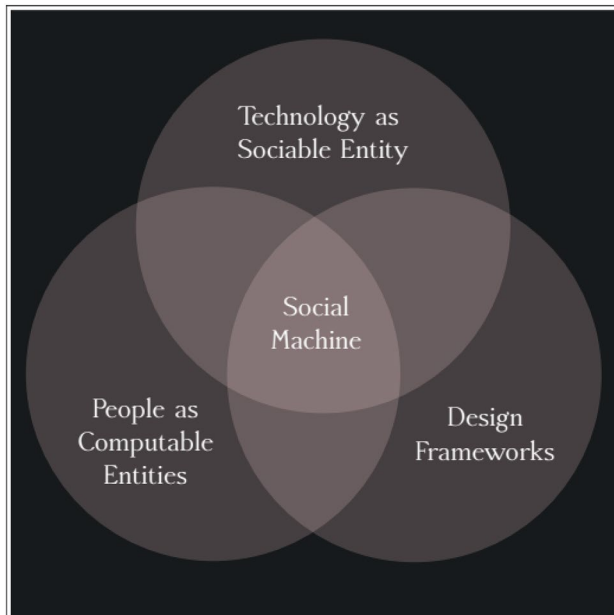


Fig. 2 Social machines are shaped by human behavior, technological implementations and design frameworks

business models, often prioritising interactions that promote efficient advertisement placement instead of the optimal interaction between users (Murray-Rust et al. 2015).

Studying social machines from a holistic framework such as cybernetics becomes even more important given the nature of contemporary human-algorithmic ecosystems. In the era of ubiquitous computing, individuals and computers are integrated into systems of circular communication and control. Individuals are constantly enhanced by algorithms integrated with most technological artifacts, be those navigation tools, social network platforms, or search engines. This pervasive, persistent, invisible and continuous existence of algorithmic applications in every aspect of human life (Roush 2005; Meira et al. 2011) distorts the classical limits between personal social autonomy and connectivity (Aakhus 2017), violates assumptions of classical causality and generates a networked space–time (Murray-Rust et al. 2015) in which the role of operator and operand are constantly exchanged between humans and algorithms.

3.3 Social machines and cybernetics: examples

A piercing example of this transcendental process where technology is no longer an aiding tool for humans but where humans also become an aiding tool of technology is social computing. The efficiency of contemporary data-intensive algorithms is mainly based on the quality of input data, which should reflect clearly, in a detailed and unbiased way, every aspect of social behaviour. Thus, humans transform themselves into datafied artifacts, offering every aspect of

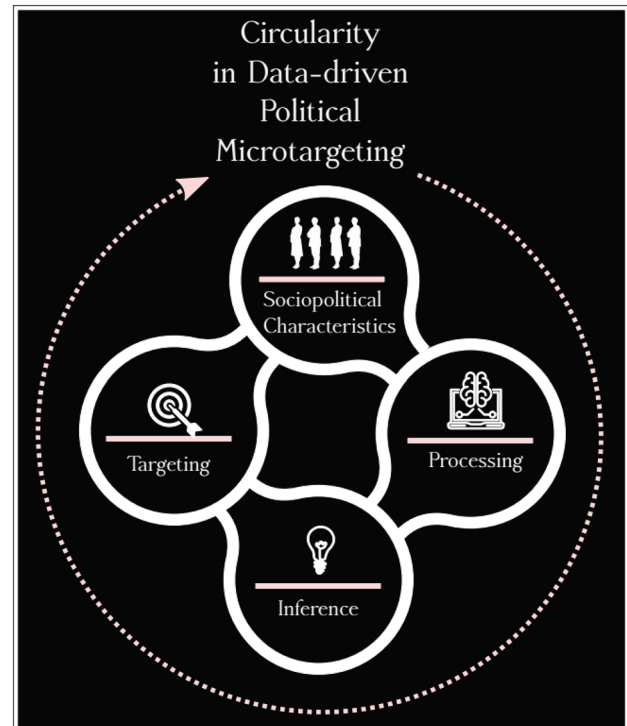


Fig. 3 Circular loop in data-drive political microtargeting, with social and computational mechanisms transcending their materiality and classical notions of causality becoming inapplicable

their lives to algorithms to optimize the latter's function. This can not only be seen in the rise of social machines largely based on crowd sourcing (Martin and Pease 2013), where people actually do the creative work and show computers what to learn and how (Berners-Lee and Fischetti 2001; Hendler and Mulvehill 2016), but also in the emergence of new structures in social and political conduct. For instance, state-of-the-art campaigning, which can be analyzed through the framework of political machines, is based on the generation of data-intensive models about the electorate and the extraction of information from them about voter interests and behaviors, which are used for the generation of ads and adapting parties' profiles (Hersh 2015; Kreiss 2016). Thus, the electorate is transformed into data for the algorithms, which then provide specific inferences to political actors, which are then transformed into actions that influence the electorate, generating a circular loop containing social and computational mechanisms, in which the notions of cause and effect become inapplicable (Fig. 3).

What remains constant in such human-algorithmic ecosystems is not the materiality of humans and algorithms, where computability and sociability become interchangeable (Murray-Rust and Robertson 2015), but the system behaviours formed by communication processes, which are dependent on how humans and algorithms interact and

influence each other (Hall et al. 2008; Katz 2017). The participants of these systems wayfare in time–space, generating dynamic meshworks of interactions, extracting information, and adapting their behavior (Murray-Rust et al. 2015), often generating fabrics of sociability and memory (De Roure et al. 2015). For example, a dating app’s emergent community is dependent both on the users’ behaviour and generated data on the platform and the ability of its algorithm to match people according to their attitudes. Similarly, the deployment of an ADM system for recidivism purposes is only feasible if it is able to remember and retrieve people’s general behaviour based on the data it was trained on.

4 Social machines, influence and politics

In social machines, individuals influence algorithms and algorithms in turn influence human behaviour. Therefore, any framework that aims to study politics should have a clear definition of what is political and what is not. Nevertheless, political theorists hold different conceptions of politics and their appearance in social interactions. Concepts such as power, influence, the social and the political are used to describe different aspects of society. For example, Laswell states that the study of politics denotes the study of influence and of the influential (Laswell 2018), referring primarily to how social groups and individuals interact and shape policy decisions. In this conduct, influence is the status which potential individuals or groups hold to achieve their purposes (Laswell and Kaplan 2013), while power is a coercive form of influence that negatively impacts the influenced. Habermas (2011), from his perspective, claims that politics is the struggle for and the exercise of power, while the political emerges as the symbolic representation and collective self-understanding of a community that reflexively deals with forms of social integration. Arendt (1972), holding a collective notion of power, states that power emerges out of the capacity to act in concert for a public-political purpose. She also states that the social is what belonged to the sphere of the private, and the past few centuries got alienated by economic and technological processes. In contrast, Foucault (1990) argues that power is everywhere, emerging between any type of social interaction, even within the family, being not external but immanent in all forms of social processes, such as economic processes, knowledge relationships and sexual relations (Foucault 1990).

4.1 Influence as a meta-concept of politics

Given the various definitions of what is and is not political, I create a conception of politics that can be useful in the study of social machines. First, technology in social machines has also a social role, thus it should be excluded neither from the

social nor the political function of the systems. Furthermore, I agree with Foucault that power is everywhere, and in our case, it can emerge also at the interaction of humans with technology. Nevertheless, I treat power as a coercive form of influence, hence a subcategory of it, like Laswell and Kaplan do (Laswell and Kaplan 2013). Influence is something that constantly emerges in social machines, regardless of the nature of the participant, be that an algorithmic recommendation a user sees or a politician targeting the electorate with personalized messages. Therefore, influence functions as a meta-concept that incorporates processes that take part at the individual level, at the social level, at the institutionalized political level and at the sociotechnological level. Regardless of the form of influence and the level of its appearance, the fact that the meta-concept encompasses different instances of power and influence makes it possible to compare and understand the complex patterns within the ecosystem, that may otherwise seem unrelated to each other.

Influence often emerges in the social domain, when individuals want to fulfill their economic, physiological, or socialization needs. Nevertheless, such processes always have a political dimension, since the way influence changes participants’ behavior has an impact on the organization, values, hierarchies and outcomes of a socioalgorithmic ecosystem. For example, the feeling a user interface color gives to an individual can impact how they evaluate a political message, leading to a chain of alterations that can have unforeseen political effects. Therefore, any social interaction that includes an influence process automatically has also a potential political component. This political component not only includes institutionalized politics, organized group behaviour, but also attitudes and behaviours of individuals that shape the rights, obligations, possibilities and boundaries of an individual or social group in the society.

4.2 From social to political machines

Since this study wants to provide a framework that assists in the understanding of all these instances of politics in socio-algorithmic ecosystems, I classify influence processes in social machines based on their type. Because every interaction in social machines has an immanent political dimension, I refer to social machines from now on as political machines. I do so to introduce a normative perspective, which states that no interaction in social machines is neutral. On the contrary, it has the potential to affect individuals and the society in unforeseen ways, therefore, their political aspect should be always taken into consideration, both during their design and their integration into the society. The political machines framework wants to exploit the advantages provided by social machine cybernetics, connect influence processes that might seem unrelated, and provide a toolkit that advances

the understanding of political phenomena shaped by humans and technology in a simple yet efficient way.

5 Political machines: a framework

Political machines include uncountable influence processes because individuals and technology are constantly interacting in them. For example, a recommendation system changes the behavior of the users on a platform, while a user's behaviour changes the system's recommendations recursively. Similarly, how a social network is designed, what its purpose is and the interaction possibilities of users, change

both user behaviour and the platform's algorithmic design. A commercial or political actor deploys an ADM system for their decisions to be influenced by the model's results. It is clear, therefore, that influence processes are constantly at work, can be of various types and can concern different participants. To understand such interactions, the political machines framework provides a five-point classification of influence processes.

Within political machines, five main categories of influence take place: A. symbolic influence, B. political conduct, C. algorithmic influence, D. design, and E. regulatory influence. Each of the above categories differently contributes to the formation and identity of political machines, how the

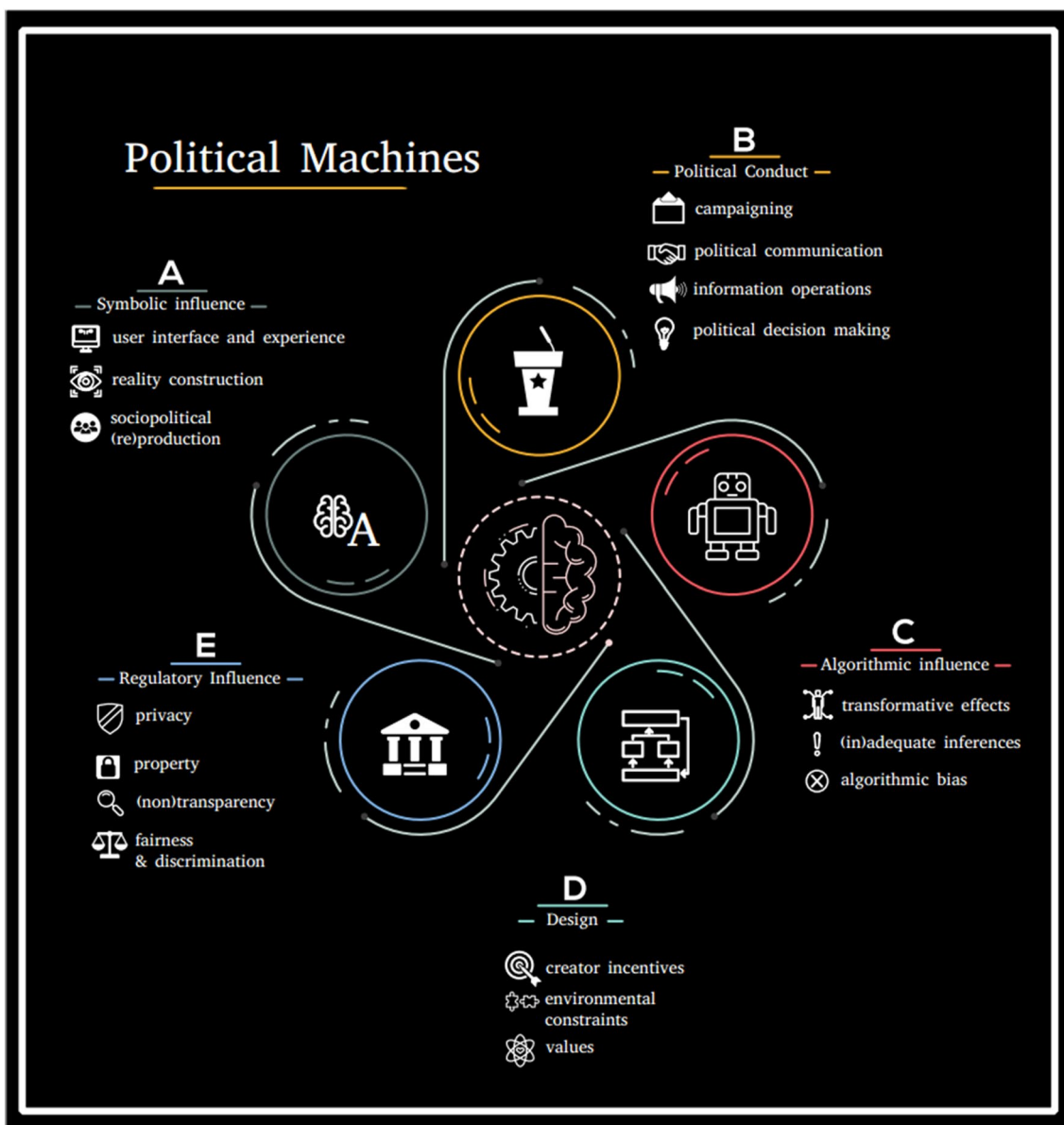


Fig. 4 The five main categories of influence in political machines

system components interact and change, and how systems reach their equilibrium (Fig. 4). In the following, I explain each of the above categories of influence.

5.1 Symbolic influence

The most basic form of influence in political machines is bound by human cognition and is of symbolic nature. The individual, to either perceive the world or to explain and communicate it, deploys symbols (Mead 1934). Human language and thought consists of words which are nothing other than symbols composed by signifiers and signifieds (De Saussure 2011; Vygotski 2012). These symbols carry with them social conditions and meanings, informing the individual about the world and influencing their behaviour. For example, the explicit inclusion of a text for opting into a platform's terms and conditions has the potential to change the decision of a user using that service. Similarly, making a list of available genders for a user to select or not functions as a proxy of social power, social group (in)visibility and illustrates existing social inequalities (Van Dijk 2001; Fairclough 2013). This also happens when users converse on platforms, even about non-political issues, with the generated text and discussions revealing and reproducing dominant social group attitudes and perceptions (Bourdieu 1979), which are then passed on to the reader.

Symbolic influence does not only appear in the context of language. Non-verbalised information, in the form of stimuli, such as seeing shapes or colours, can also be responsible for influencing an individual. Such information is stored in human memory as mental representations or information schemes (Piaget 1947), which are reactivated, retrieved and deployed depending on new incoming information. Therefore, the appearance and structure of a user interface (UI) and the linked user experience (UX) can always influence participants in political machines, changing their behavior. For example, a platform's UI color can influence how much time a user spends with a service (Shneiderman and Plaisant 2010), or how much and in which way they would interact with it (Benyon 2014). Symbolic influence encompasses such processes, focusing on the power that symbols have in shaping social reality and behaviour. Of course, which symbols will appear in a political machine is largely driven by the incentives of political machines owners and designers, and such decisions belong to the dimension of design, as it will be analyzed next.

Because implicit or explicit symbols always appear in political machines, symbolic influence is the most subtle and penetrating type of the influence. It is always there, but the full extent of its impact is practically impossible to quantify. Nevertheless, in specific cases with appropriate experimental design, researchers can investigate and understand the properties of symbolic influence (King et al. 2017).

5.2 Political conduct

The most straightforward way in which politics appears in political machines is when politicians and political actors use them as means of improving their status and increasing their power, or when democratic processes explicitly take place in them. Therefore, what is here referred to as political conduct encompasses any social group, individual or institutionalized actions that explicitly and consciously have a political motive. This includes cases where participants on political machines are actively seeking to transform society and influence the existing hierarchical and power structures.

Political conduct appears often in online social networks and ADM systems. Although most prominent social media platforms were not designed to foster political discussions, nowadays they serve as central spaces for political exchange, campaigning and communication. Users utilize platforms to comment on civic and political issues, externalize their political ideologies and form online groups of political action (Rainie et al. 2012; Gustafsson 2012). This ample space for political interactions generated hopes and promises for a more diverse, open and democratic political discourse (Loader and Mercea 2012). Social media platforms were treated as space for more autonomous political acting (Fenton and Barassi 2011), which could contribute to the diffusion of voices that were systematically repressed by authoritarian regimes and power structures (Joseph 2012; Rentschler 2014). These expectations were largely generated because social media can be a space for information, connection, mobilization, deliberation and diversity (Zuckerman 2019).

Nevertheless, social media as a space for political conduct became mainstream and got exploited by various political actors. Contemporary politicians and political parties maintain pages and profiles on social media platforms to interact with the electorate. Simultaneously, they deploy large-scale political campaigns to influence public opinion, especially in the form of personalised advertising (Zuiderveen Borgesius et al. 2018; Medina Serrano et al. 2020). The openness of social media is exploited by other actors as well, with automated, fake, and militant accounts spreading misinformation (Ratkiewicz et al. 2011; Howard and Woolley 2016).

All the cases above constitute social media as highly complex political. Even the political processes taking place explicitly on them are of different forms, with multiple participants using the services towards their specific goals. In this context, many debates take place to investigate whether social media actually contributes to the democratisation of society or has a negative political impact (Effing et al. 2011; Zuckerman 2014; Gorham 2020; Bennett 2012).

Besides the political conduct taking place in social media, a variety of ADM systems are deployed for political purposes. Political parties hold large databases containing

demographic and personal information, on which they train data-intensive models used then for decision making in political campaigns (Kreiss 2016; Hersh 2015). In many cases, computer vision tools assist police in detecting suspects (Idrees et al. 2018), while there are many companies developing systems that quantify the propensity of individuals to commit crimes. These systems are often used by criminal justice (Dressel and Farid 2018). Especially in the legislature, ADM systems are increasingly developed and deployed for automating legal evaluations and detecting violations (Ashley 2017).

It is clear from the above that technology and political conduct intersect both on various social machines, with political actors using technology towards their ends. Given the complexity of interactions, many open questions exist about the direct politicization of technology and its recursive impact on the political processes taking place.

5.3 Algorithmic influence

One of the most important questions related to political machines is how algorithms directly influence individuals. Algorithmic influence is a cardinal part of many political machines because services and political actors explicitly deploy algorithms for automating processes, affecting individuals and society. Algorithmic influence encompasses processes caused by the mathematical structure, predictions and inferences of an algorithm. Of course, under which criteria such features will be implemented are strongly influenced by the goals, needs, and values of algorithms' owners. These decisions belong to the category of design, which will be analyzed next.

Algorithmic influence of the society takes place in both political and non-political settings, with ubiquitous computing covering every aspect of socialization. From google maps to online content suggestion, human behaviour is constantly reshaped by algorithmic implementations. For example, on social media, platform designers deploy algorithms to 1. suggest personalized content to users, 2. to place targeted advertisement, and 3. to filter and review the contents generated by users. All three algorithmic implementations have the potential to change human behavior in different ways.

By selecting which contents are going to be visible to a user's news feed, an algorithm leads to reality tailoring (Just and Latzer 2017). What a user perceives about the world changes in respect to the selected pieces of information, leading to an algorithm-mediated subjective knowledge. That knowledge is then transformed into actions, with users forming opinions about the world and actively behaving according to them in the online and offline world. In this context, it has been largely hypothesized that algorithms

can lead to filter-bubble phenomena (Pariser 2011). Filter bubbles are segregated opinion clusters formed by the algorithms, in which users only come into contact with conforming opinions but not to opposing ones, a social setting that can easily lead to opinion polarization.

Even if this content curation does not lead to polarization, it always introduces a bias, because the algorithm-mediated reality is dependent on an algorithm's structure and the related input data. This leads to the emergence of data politics (Ruppert et al. 2017) that concern themselves with how algorithms function (Seaver 2019), what biases they introduce (Lazer 2015; Bozdog 2013), and how they influence the individuals and social groups (Taylor 2017; Beer 2017). Data politics does not restrict itself to recommendation algorithms on social media, but also includes the platforms' services for personalized advertisement in the form of microtargeting (Kreiss 2016; Hersh 2015). These opaque algorithms place advertisements to users according to demographic and behavioral criteria with the aim of efficiently influencing user behaviour. Microtargeting is the state-of-the-art technique in political campaigning, while the platforms' business models largely depend on convincing commercial companies and political actors to rent these services for advertising.

Another dimension of algorithmic influence on social media is related to content filtering algorithms. Companies largely use automated processes that scan uploaded images, videos, and text and search for contents that violate the platforms' terms and conditions. These algorithms, therefore, decide what is allowed to become part of the open discourses and what not, how freedom of speech is constituted on the platforms, and coordinate the development of user behavior.

Algorithmic influence is equally present in other types of social machines, such as ADM algorithms. ADM systems are largely used for risk assessment and warning (Mosier and Skitka 1996), for automating tasks such as image recognition, speech understanding, medical consulting, and predictive policing (Larus et al. 2018; Ensign et al. 2017; Dressel and Farid 2018). ADM systems result in bidirectional algorithmic influence. First, they influence the behavior of the users who deploy the models, because they generate knowledge that is exploited in multiple decision-making processes. Second, in the case that algorithms are also making decisions about individuals and social groups, their decisions influence these groups. For example, a hiring algorithm does not only influence the company by suggesting a candidate but also influences the candidates themselves, deciding who will get a job or not (Ajunwa et al. 2016). An algorithm that recommends treatment type for patients to a doctor not only helps the doctor in making the optimal decision but also chooses whether patients should get operated or not and how long their recovery period will be, etc. In such cases, algorithmic inferences often raise epistemic

concerns regarding their predictive power, accuracy, and their general adequacy to provide reliable evidence for a decision (Mittelstadt et al. 2016). Since algorithmic implementations are opaque, do not provide sufficient evidence for their inferences, and sometimes do not provide “hard” answers, the application of algorithms in contexts such as the above remains questionable.

Furthermore, because algorithmic impact is dependent on multiple parameters such as input data, social structures, and designers’ choices, there is an open question about how and under what conditions algorithms remain neutral media in decision making and content management. Multiple cases show that algorithmic implementations discriminate individuals and social groups, resulting in unfair decisions and politically influencing public opinion (Introna and Nissenbaum 2000; Lustig et al. 2016; Friedler et al. 2016). Therefore, algorithms have the potential not only to change human behavior but also to perform these changes in an asymmetric way, which often violates ethical norms and social expectations. Given the above, algorithmic influence is of high complexity and dimensionality; it is a challenge for researchers to understand it and for political actors to regulate it.

5.4 Design

The fourth dimension of political machines is related to the systems’ design. How symbolic influence, algorithmic influence, and political conduct take place, depends on the structure of the political machines, which are largely given by the design of their components. Each component of a political machine takes its final form given the designers’ objectives and existing environmental constraints (Newell and Simon 1972). This final form is going to contribute to the resulting equilibrium in a political machine. For example, the design principles of a social credit system influences the behavior of citizens in a society, setting the peoples’ feasible action space, and forming their social goals (Pasquale 2015; Engelmann et al. 2019). Similarly, a social media recommendation system suggests contents to the users in a way that aligns with the company’s business model goals.

Design constraints also have a huge impact on the formation of political machines. Hardware or software limitations can result in discriminative model predictions (Sandvig et al. 2014) even if that was not part of the designers’ intentions. The ability of a model in predictive medicine to make good decisions is dependent on the available data which, given privacy issues, might be scarce and, therefore, lead to a model deployment with lower predictive ability.

Besides the designers’ goals and environmental constraints, a parameter that strongly influences the formation of political machines are design ethics. The decision of tech companies to gather data about users’ interests, traits, demographic and behavioral information, and exploit them into

developing better algorithms is always dependent (Hitlin and Rainie 2019) on the owners’ perception of what is ethical, what is necessary for achieving their goals, and what is allowed by the state. The fact that companies do not disclose how their systems work, maintaining a high level of opacity in every aspect of the models development and deployment, is a design property that obstructs the understanding of the systems and determines the accountability and transparency (Sandvig et al. 2014) between the state, users, and systems owners. Especially in cases of auditing algorithms and trying to trace their potential discriminatory impact or political influence, such design properties obstruct researchers from interpreting phenomena and from good societal governance (Barocas et al. 2013).

Issues like the above raise questions about how to ideally design political machines that serve the society in an optimal way, given that many technological ecosystems are driven by financial incentives (Langlois and Elmer 2013), having unknown transformative effects on politics and society. For example, although political communication largely takes place on social media, these are not public, nor do they try to always remain politically impartial (Engelmann et al. 2018; Leskovec et al. 2010). The same applies for ADM systems that are deployed either by the state or are of high social value (Lepri et al. 2018), which provide inferences in terms of mathematical probabilities (Ananny 2016). Arguing, justifying and legitimizing an action based on a probability is can be problematic, as a probability scholastically evaluates a situation and does not deterministically result in an inference.

The above cases are only a few examples of how design values, creators’ incentives and environmental constraints can influence the formation of political machines. The analysis of each political machine can reveal multiple design properties that constitute the participants’ interactions on them. Therefore, a detailed analysis is necessary for an exact political evaluation of them.

5.5 Regulatory influence

The last form of influence in political machines is related to regulatory frameworks. Because political machines are embedded within a society, and since societal function is controlled by institutionalized processes (Castoriadis 1997), political and legal structures set the space in which political machines can function. In algorithmic applications, legislature decides how these systems should be deployed, and how the interests of the designers and the public can be protected. Main regulatory issues for algorithmic applications are related to data property and privacy, algorithmic opacity, and discrimination of social groups. In the following, I provide an overview of these topics.

- *Data property and privacy* One of the main reasons for the current intensive application of algorithms is the datafication that has taken place since the beginning of the third millennium. The vast amount of created data related to human behavior can be exploited and used to multiple ends, with new business models being born unstoppably. Data is generated, collected, processed and combined for decision making, political and commercial acting, raising questions about whom this data actually belongs to, what kind of rights a data collector has, as well as whether the collection and data processing violates individuals privacy rights (Bertot et al. 2010). For answering such questions, states possess various regulatory frameworks that define what is allowed and what is not (U-Directive 2016).
- *Algorithmic opacity* One of the main designers' rights in algorithmic implementations is their legal protection in not disclosing their models' inputs, structures, and outputs. This is because a developed model can provide better market opportunities to its owner, therefore, its features can remain secret under the fear of competition. Nevertheless, the resulting algorithmic opacity obstructs the auditing and understanding of such systems, especially when it comes to algorithmic impacts that violate the law (Burrell 2016).
- *Discrimination and freedom of expression* Legal frameworks interfere with discriminative political machines in two ways. First, as already discussed, algorithmic implementations might result in discriminatory decisions against individuals and social groups. Especially for ADM applications, practice proves that such events can happen frequently, raising questions about the extent to which existing legislations are violated regarding protected social groups, and individual rights and freedoms (Zarsky 2016). Second, on social media and other online platforms, algorithms are deployed for content filtering. This happens for two reasons: First, platforms remove content that contains harmful and discriminative speech or violates legislation for other reasons. Second, platforms want to protect their service function and thus remove content that does not comply with their imperatives. In this process, questions arise about (1) when does a specific content violate legislation, (2) how is individual freedom of expression defined and where does society set its limit, (3) who is legally accountable for contents that were wrongly unfiltered and contents that were mistakenly filtered, and (4) how free should companies be in choosing what to filter or not given their financial incentives.

Overall, algorithmic implementations in political machines remain largely unregulated. Given this, a lot of discussions take place around algorithms and their definition

(Ziewitz 2016; Seaver 2017), their current and ideal functions (Bertot et al. 2012), how regulations could prevent biases inserted by algorithmic applications (Introna and Wood 2004), as well as who should be accountable in cases of misconduct (Barocas et al. 2013; Diakopoulos 2014). Regulation, therefore, emerges as one of the most crucial categories of influence, because it has the potential to transform the very nature of political machines.

5.6 Illustration of the framework

The above categories of influence in political machines are neither static nor independent. They interact constantly, dynamically transforming the systems' states and shaping how individuals and the society will behave. Each form of influence within a category not only reforms political machines but has its feasible space defined by the influence processes of other categories. In the following, I present two case studies of political importance in political machines and illustrate how influence processes of various categories interacted with each other. I demonstrate how the framework is able to reduce political complexity within political machines, to contribute to the understanding of explicit and implicit systemic changes, and to guide researchers, policy-makers, and designers in evaluating how interventions could shape ecosystems (RQ2). I do so by exploiting graphs as means for connecting influence processes of different nature, as well as using tabular representations to associate events under a cybernetic framework.

5.7 Reducing exposure to alt-right content

An example of how specific changes in political machines can influence multiple components both directly and indirectly was the decision of YouTube in 2019 to minimize the exposure of users to alt-right political content. This decision was drawn partly based off scientific evidence showing strong radicalization patterns on the platform (Ribeiro et al. 2020), with users progressively moving from the consumption of right-wing moderate content to far-right ideological content. This content decision was associated to the platform's design values however its operationalization took place by excluding the specific content from platform's recommendation algorithms. This algorithmic alteration indeed altered user behaviour, reducing the popularity of such content (Buntain et al. 2020). Nevertheless, researchers showed another indirect effect of this decision. The change of the algorithm's function not only affected YouTube, but a similar decline in the consumption of such content was detected both on Reddit and Twitter (Buntain et al. 2020). This shows that the change in political communication on one platform can alter the political conduct on other social media platforms as well. Figure 5 presents

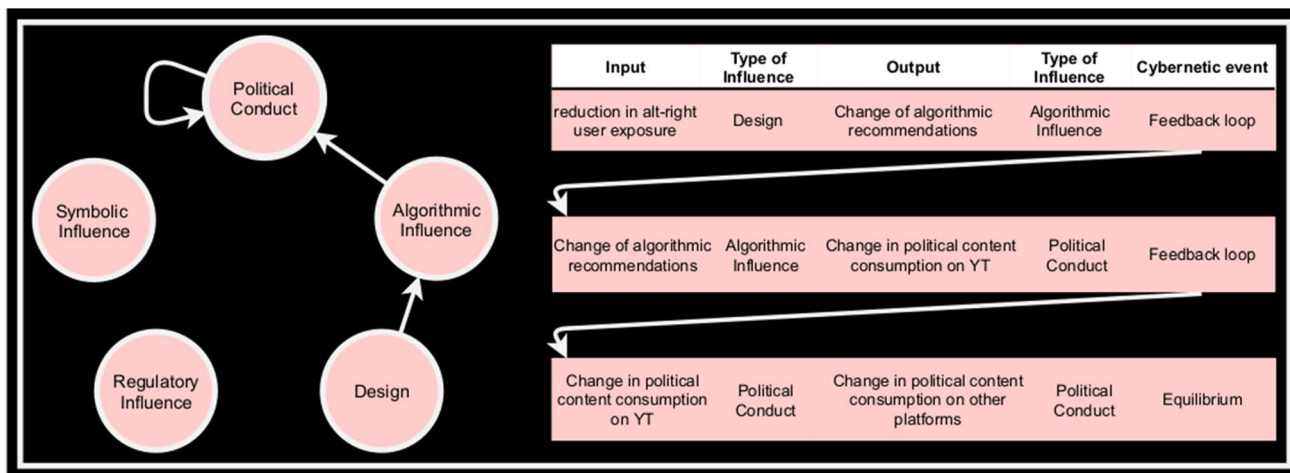


Fig. 5 Analyzing the decision of YouTube to reduce user exposure to alt-right content. The graph on the left shows the interconnectedness of influence processes. The table on the right illustrates the related cybernetic analysis

the interconnectedness of influence processes when studying the total social media ecosystem as a political machine. A cascade of influence processes took place in three different domains: design, algorithmic influence, and political conduct. From a cybernetic perspective, the change in design values lead to multiple feedback loops, meaning that outputs of systemic processes recursively became inputs, having further effects. In the end, the system reached an equilibrium which was associated with lower consumption of alt-right content on three different platforms.

5.8 Data-driven political microtargeting, Facebook, and the GDPR

Data-driven political microtargeting as a campaign strategy exploits algorithmic-decision making systems to generate inferences about the electorate and targets it with personalized advertisements (Hersh 2015). This campaign practice caused platforms to adapt their ad targeting systems, to attract customers. For example, Facebook offers the option to target individuals based on their inferred political preferences, a feature that their algorithm exploits to decide who is going to see the political content and who is not (Analytics, n.d. 2019). Nevertheless, this option is available only in the US, since the regulatory framework allows it. In contrast, such a platform service is not feasible in Europe, since the European General Data Protection Regulation (GDPR) explicitly defines the limits and possibilities of using data for political purposes (U-Directive 2016). Figure 6 analyzes the above processes using the political machines framework. The graph shows that political campaigns influences Facebook design, who in turn adapts their algorithmic structure, altering who will be targeted. Furthermore, political campaigns also adapt their political conduct, since they use

Facebook ad services to reach the electorate. Regulatory frameworks, therefore, function as systemic constraints, because they define the feasible space of how to place political ads for each country. From a cybernetic perspective, the above interactions reach an equilibrium in which political campaigns in the US and in Europe can use Facebook ad services in different ways.

5.9 Knowledge extraction and potential of the framework

The above examples illustrate how interconnected influence processes are in political machines. The framework serves as a tool for evaluating such interactions in a more structured way. Processes can be categorized by their respective influence categories, and their effects can be connected to the other categories, reducing the systems’ complexity and tracing direct and indirect relations. For example, it is visible from the two case studies that the way in which algorithms influence users on social platforms is dependent on design decisions of platform owners. Such evidence can be used as an argument for supporting accountability claims related to unjust and problematic algorithmic inferences. Of course, it is necessary to complement any extracted knowledge through further scientific theories that can transform evidence into structured arguments.

The political machines framework may not only be applied for evaluating complex processes of influence taking place in political machines, but also for the purposes of political machine design. By intervening in a system and holding everything else equal, scientists can uncover how a single change in a political machine can influence multiple processes, resulting in impact assessment and quantification. Scientists can assess how a new platform feature

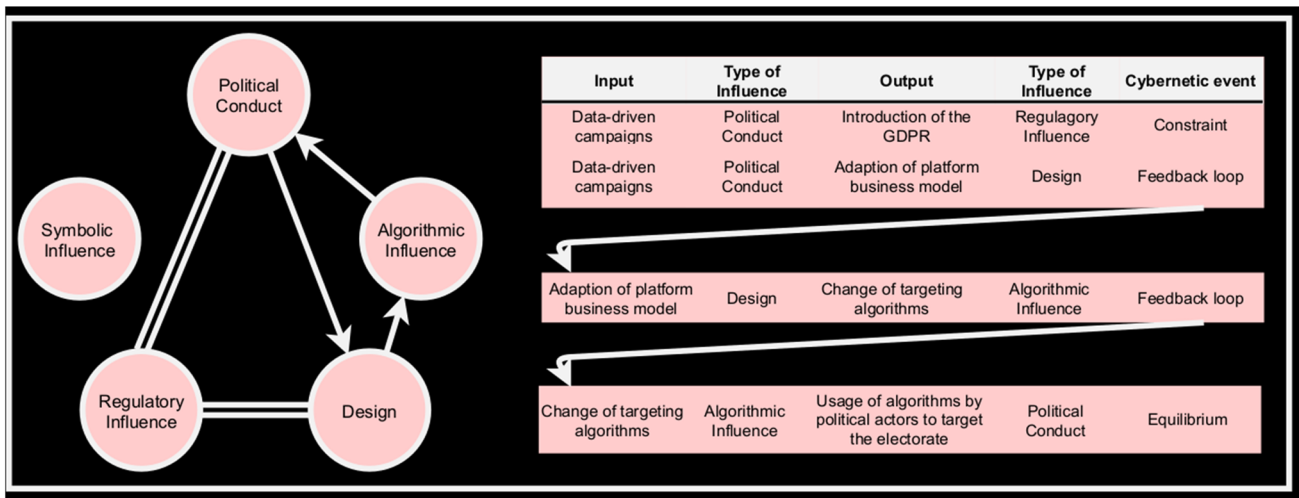


Fig. 6 Analyzing data-driven political microtargeting from the perspective of Facebook. Regulation functions as a systemic constraint, defining what is feasible. Political conduct of political campaigns,

Facebook design values, and algorithmic systems form a circular loop of communication and control

might influence political conduct, or how a new regulation changes the structure of a political machine. This is of high importance in an era where socio-algorithmic ecosystems are largely unregulated. Designers maintain a high degree of opacity around their systems which translates to a lower degree of accountability. Political machines as a framework can support researchers in answering questions such as the above, and for revealing and understanding the uncountable influence processes existing in the ecosystems.

6 Implications, limitations and future work

The introduced framework successfully answers RQ1: it provides a way to analyze and classify political processes in political machines, by adopting a systemic perspective. In this way, it fills an important scientific gap, since no attempts have been made to analyze politics on socio-algorithmic ecosystems from a holistic point of view. The framework achieves that by investigating the interplay of political processes that are usually not analyzed or thought to belong together. It does this, by detaching humans and technology from their materiality, and focusing on what each of them does.

The analysis also answers RQ2: The framework provides a structure that can guide researchers in understanding, designing, and intervening in socio-algorithmic ecosystems. Using case studies, it was demonstrated that the framework can successfully evaluate occurring influence processes, even in cases where systems are of high complexity and a single change can have unforeseeable effects. The analyzed case studies described how inputs

shift political machines' equilibria, and can be used as a tool for complementing scientific theories in knowledge extraction. Furthermore, since socio-algorithmic ecosystems constantly face important ethical and political challenges, the framework can be used to semantically plan how potential interventions, either in the regulation, or in the political machine itself, might change systems' dynamics.

Since the framework is applied for the conceptual understanding of politics in socio-algorithmic ecosystems, it also comes with specific limitations. First, there is a need for a systemic analysis of existing political machines and their immanent political processes, to trace regularities of influence that take place. In this way, specific events can be linked to each category of influence, and additional and graspable knowledge can be generated about the nature of politics for each type of political machine. Second, although the framework was applied in this study to understand past interactions within political machines, further empirical work should be performed that uses the framework in ongoing interventions, to verify its ability to predict and guide researchers, policymakers, and platform designers in their work.

Despite these limitations, even in the short examples demonstrated above, additional knowledge was generated regarding political processes in socio-algorithmic ecosystems. In most cases of political machines, there exists a political equilibrium, where platforms' owners and algorithmic designers have the most control in the systems' function. This systemic feature initiates a discussion about how political machines should be and how they should be designed. Most algorithmic implementations today are

part of the commercial sector, with states having marginal control over them and regulators facing serious challenges. Furthermore, individuals and social groups are the most passive participants in the system, usually taking either the role of the consumer, or being projected into datafied artifacts. From a normative perspective, society should reflect on the meaning of these roles, and re-imagine the future of socio-algorithmic ecosystems.

Centering civic interest, and the idea that technology should serve individuals and the society in a way that ensures equality, justice, political freedom and social inclusiveness, the study of political machines should be extended. Researchers should not only describe how political machines function, but also define principles, frameworks, and constraints that can lead to the creation of socio-algorithmic ecosystems that serve the public interest. The design of civic machines prevails as a necessity in an environment where technological and algorithmic implementations influence society in unexpected ways, transforming the political essence of society. This study made a first step in that direction, by defining political machines and introducing a framework for analyzing politics in socio-algorithmic ecosystems. There is an endless space for further scientific investigation, and the new knowledge can be used to create political machines, by the society and for society.

Acknowledgements This work was supported by the Center for Information Technology Policy at Princeton University and the Princeton University Library Open Access Fund. The author would like to thank Arwa Michelle Mboya and the three anonymous reviewers for their valuable feedback on the manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aakhus M (2017) Understanding information and communication technology and infrastructure in everyday life: struggling with communication-at-a-distance. In: *Machines that become us*. Routledge, pp 27–42
- Ahlers D, Driscoll P, Löfström E, Krogstie J, Wyckmans A (2016) Understanding smart cities as social machines. In: *Proceedings of the 25th international conference companion on World Wide Web*, pp 759–764
- Ajunwa I, Friedler S, Scheidegger CE, Venkatasubramanian S (2016) Hiring by algorithm: predicting and preventing disparate impact. Available at SSRN
- Analytics, NewsWhip. n.d. (2019) The 2019 guide to Facebook publishing. *Social media analytics*. http://go.newswhip.com/2019_03FacebookPublishing_LP.html
- Ananny M (2016) Toward an ethics of algorithms: convening, observation, probability, and timeliness. *Sci Technol Hum Values* 41(1):93–117
- Arendt H (1972) *Crises of the republic: lying in politics, civil disobedience on violence, thoughts on politics, and revolution*, vol 219. Houghton Mifflin Harcourt, Boston
- Ashby WR (1957) *An introduction to cybernetics*. Chapman & Hall Ltd, London
- Ashley KD (2017) *Artificial intelligence and legal analytics: new tools for law practice in the digital age*. Cambridge University Press, Cambridge
- Bakshy E, Messing S, Adamic LA (2015) Exposure to ideologically diverse news and opinion on Facebook. *Science* 348(6239):1130–1132
- Barberá P, Jost JT, Nagler J, Tucker JA, Bonneau R (2015) Tweeting from left to right: is online political communication more than an echo chamber? *Psychol Sci* 26(10):1531–1542
- Barocas S, Hardt M, Narayanan A (2017) *Fairness in machine learning*. NIPS Tutor 1:2
- Barocas S, Sophie H, Malte Z (2013) *Governing algorithms: a provocation piece*. Available at SSRN 2245322
- Bateson G (2006) A theory of play and fantasy. In: *The game design reader: a rules of play anthology*, pp 314–328
- Beer D (2017) *The social power of algorithms*. Taylor & Francis
- Bennett WL (2012) The personalization of politics: political identity, social media, and changing patterns of participation. *Ann Am Acad Pol Soc Sci* 644(1):20–39
- Benyon D (2014) *Designing interactive systems: a comprehensive guide to HCI, UX and interaction design*. Pearson Edinburgh, Edinburgh
- Berners-Lee T, Fischetti M (2001) *Weaving the web: the original design and ultimate destiny of the World Wide Web by its inventor*. DIANE Publishing Company, Darby
- Bertot JC, Jaeger PT, Hansen D (2012) The impact of policies on government social media usage: issues, challenges, and recommendations. *Gov Inf Q* 29(1):30–40
- Bertot JC, Jaeger PT, Munson S, Glaisyer T (2010) Social media technology and government transparency. *Computer* 43(11):53–59
- Bird S, Kenthapadi K, Kiciman E, Mitchell M (2019) Fairness-aware machine learning: practical challenges and lessons learned. In: *Proceedings of the twelfth ACM international conference on web search and data mining*, pp 834–835
- Bolukbasi T, Chang KW, Zou J, Saligrama V, Kalai A (2016) Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In: *Advances in neural information processing systems*, pp 4349–4357
- Bourdieu P (1979) Symbolic power. *Crit Anthropol* 4(13–14):77–85
- Bozdag E (2013) Bias in algorithmic filtering and personalization. *Ethics Inf Technol* 15(3):209–227
- Buntain C, Bonneau R, Nagler J, Tucker JA (2020) YouTube recommendations and effects on sharing across online social platforms. *arXiv Preprint arXiv: 2003.00970*
- Buolamwini J, Gebru T (2018) Gender shades: intersectional accuracy disparities in commercial gender classification. In: *Conference on fairness, accountability and transparency*, pp 77–91
- Buregio V, Meira S, Rosa N (2013) Social machines: a unified paradigm to describe social web-oriented systems. In: *Proceedings of the 22nd international conference on World Wide Web*, pp 885–890

- Burrell J (2016) How the machine ‘thinks’: understanding opacity in machine learning algorithms. *Big Data Soc* 3(1):205395171562251
- Castoriadis C (1997) *The Imaginary Institution of Society*. MIT Press, Cambridge, MA
- Chouldechova A, Roth A (2018) The frontiers of fairness in machine learning. arXiv Preprint arXiv: 1810.08810
- Cristianini N, Scantamburlo T (2019) On social machines for algorithmic regulation. *AI Soc* 35:1–18
- Dalton CM, Taylor L, Thatcher J (2016) Critical data studies: a dialog on data and space. *Big Data Soc* 3(1):2053951716648346
- De Roure D, Clare H, Kevin P, Ségolène T, Pip W (2015) Observing social machines part 2: how to observe? In: Proceedings of the ACM web science conference. ACM, pp 13
- De Saussure F, Baskin W (2011) Course in general linguistics (trans: Wade Baskin). In: Meisel P, Saussy H (eds) Columbia University Press, New York. <https://doi.org/10.7312/saus15726>
- Deutsch KW (1963) *The nerves of government: models of political communication and control*. The Free Press of Glencoe, New York, p 316
- Diakopoulos N (2014) Algorithmic accountability reporting: on the investigation of black boxes. Report, Tow Center for Digital Journalism, Columbia University
- Van Dijk TA (2001) 18 Critical discourse analysis. *Handb Discourse Anal* 33:49–71
- Dressel J, Farid H (2018) The accuracy, fairness, and limits of predicting recidivism. *Sci Adv* 4(1):eaao580
- Effing R, Hillegersberg JV, Huibers T (2011) Social media and political participation: are Facebook, Twitter and Youtube democratizing our political systems? In: International conference on electronic participation. Springer, pp 25–35
- Endres K (2016) The accuracy of microtargeted policy positions. *PS Political Sci Politics* 49(4):771–774
- Engelmann S, Grossklags J, Papakyriakopoulos O (2018) A democracy called Facebook? Participation as a privacy strategy on social media. In: Privacy technologies and policy: 6th annual privacy forum, APF 2018, Barcelona, Spain, June 13–14, 2018, revised selected papers. Springer, pp 91–108
- Engelmann S, Chen M, Fischer F, Kao CY, Grossklags J (2019) Clear sanctions, vague rewards: how china’s social credit system currently defines “Good” and “Bad” Behavior. In: Proceedings of the conference on fairness, accountability, and transparency, pp 69–78
- Ensign D, Friedler SA, Neville S, Scheidegger C, Venkatasubramanian S (2017) Runaway feedback loops in predictive policing. arXiv Preprint arXiv: 1706.0984
- Erickson BJ, Korfiatis P, Akkus Z, Kline TL (2017) Machine learning for medical imaging. *Radiographics* 37(2):505–515
- Fairclough N (2013) *Critical discourse analysis: the critical study of language*. Routledge, New York
- Faris R, Roberts H, Etling B, Bourassa N, Zuckerman E, Benkler Y (2017) *Partisanship, propaganda, and disinformation: online media and the 2016 us presidential election*, vol 6. Berkman Klein Center Research Publication, Cambridge
- Fenton N, Barassi V (2011) Alternative media and social networking sites: the politics of individuation and political participation. *Commun Rev* 14(3):179–196
- Ferrara E, Varol O, Davis C, Menczer F, Flammini A (2016) The rise of social bots. *Commun ACM* 59(7):96–104
- Von Foerster H (2007) *Understanding understanding: essays on cybernetics and cognition*. Springer Science & Business Media, Berlin
- Foucault M (1990) *The history of sexuality: an introduction*. Vintage, New York
- Friedler, SA, Carlos S, Suresh V (2016) On the (Im) possibility of fairness. arXiv Preprint arXiv: 1609.07236
- Gorham AE (2020) Anonymous’s glory. *Int J Commun* 14:19
- Gustafsson N (2012) The subtle nature of facebook politics: swedish social network site users and political participation. *New Media Soc* 14(7):1111–1127
- Habermas J (2011) The political: the rational meaning of a questionable inheritance of political theology. *Power Relig Public Sphere*:15–33
- Hall W, De Roure D, Shadbolt N (2008) The evolution of the web and implications for eResearch. *Philos Trans R Soc A Math Phys Eng Sci* 367(1890):991–1001
- Haraway D (1988) Situated knowledges: the science question in feminism and the privilege of partial perspective. *Fem Stud* 14(3):575–599
- Hendler J, Berners-Lee T (2010) From the semantic web to social machines: a research challenge for AI on the World Wide Web. *Artif Intell* 174(2):156–161
- Hendler J, Mulvehill AM (2016) *Social machines: the coming collision of artificial intelligence, social networking, and humanity*. Apress, New York
- Hersh ED (2015) *Hacking the Electorate: How Campaigns Perceive Voters*. Cambridge University Press, New York
- Hitlin P, Rainie L (2019) Facebook algorithms and personal data. *Pew Res Center*:1–22
- Howard PN, Woolley SC (2016) Political communication, computational propaganda, and autonomous agents-introduction. *Int J Commun* 10:4882–4890
- Idrees H, Shah M, Surette R (2018) Enhancing camera surveillance using computer vision: a research note. *Polic Int J* 41(2):292–307
- Iliadis A, Russo F (2016) Critical data studies: an introduction. *Big Data Soc* 3(2):2053951716674238
- Introna L, Nissenbaum H (2000) Defining the web: the politics of search engines. *Computer* 33(1):54–62
- Introna L, Wood D (2004) Picturing algorithmic surveillance: the politics of facial recognition systems. *Surveill Soc* 2(2/3):177–198
- Jonas H (1953) A critique of cybernetics. *Soc Res* 2:172–192
- Joseph S (2012) Social media, political change, and human rights. *BC Int’l & Comp L Rev* 35:145
- Just N, Latzer M (2017) Governance by algorithms: reality construction by algorithmic selection on the internet. *Media Cult Soc* 39(2):238–258
- Katz JE (2017) *Machines that become us: the social context of personal communication technology*. Routledge, New York
- King R, Churchill EF, Tan C (2017) *Designing with data: improving the user experience with a/B testing*. O’Reilly Media Inc., Newton
- Kreiss D (2016) *Prototype politics: technology-intensive campaigning and the data of democracy*. Oxford University Press, New York
- Krippendorff K (2007) The cybernetics of design and the design of cybernetics. *Kybernetes* 36(9/10):1381–1392. <https://doi.org/10.1108/03684920710827364>
- Kruikemeier S, Sezgin M, Boerman SC (2016) Political microtargeting: relationship between personalized advertising on facebook and voters’ responses. *Cyberpsychol Behav Soc Netw* 19(6):367–372
- Kusner MJ, Joshua L, Chris R, Ricardo S (2017) Counterfactual fairness. In: *Advances in neural information processing systems*, pp 4066–4076
- Langlois G, Elmer G (2013) The research politics of social media platforms. *Cult Mach* 14:1–17
- Larson DB, Chen MC, Lungren MP, Halabi SS, Stence NV, Langlotz CP (2018) Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs. *Radiology* 287(1):313–322
- Larus J, Hankin C, Carson SG, Christen M, Crafa S, Grau O, Kirchner C et al (2018) When computers decide: european recommendations on machine-learned automated decision making. ACM, New York

- Lasswell HD (2018) Politics: who gets what, when, how. Pickle Partners Publishing, Auckland
- Lasswell HD, Abraham K (2013) Power and society: a framework for political inquiry. Transaction Publishers, Piscataway
- Latour B (2005) An introduction to actor-network-theory. Reassembling the social. Oxford University Press, Oxford
- Lazer D (2015) The rise of the social algorithm. *Science* 348(6239):1090–1091
- Lepri B, Oliver N, Letouzé E, Pentland A, Vinck P (2018) Fair, transparent, and accountable algorithmic decision-making processes. *Philos Technol* 31(4):611–627
- Lepskiy V (2018) Evolution of cybernetics: philosophical and methodological analysis. *Kybernetes*, pp 249–261
- Leskovec J, Huttenlocher D, Kleinberg J (2010) Governance in social media: a case study of the Wikipedia promotion process. In: Fourth international AAAI conference on weblogs and social media
- Loader BD, Mercea D (2012) Networking democracy? Social media innovations in participatory politics: Brian d. Loader and Dan Mercea. In: *Social media and democracy*. Routledge, pp 12–21
- Luhmann N (1999) *Die Gesellschaft Der Gesellschaft*. Wiss. Buchgesellschaft, Darmstadt
- Lustig C, Pine K, Nardi B, Irani L, Lee MK, Nafus D, Sandvig C (2016) Algorithmic authority: the ethics, politics, and economics of algorithms that interpret, decide, and manage. In: *Proceedings of the 2016 chi conference extended abstracts on human factors in computing systems*. ACM, 1057–1062
- Martin U, Pease A (2013) Mathematical practice, crowdsourcing, and social machines. In: *International conference on intelligent computer mathematics*. Springer, 98–119
- Mayer-Schönberger V, Cukier K (2013) *Big data: a revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt, Boston
- Mead GH (1934) *Mind, self and society*, vol 111. Chicago University of Chicago Press, Chicago
- Mead M (1968) Cybernetics of cybernetics. In: Foerster H von, White J, Peterson L, Russell J (eds) *Purposive systems*. Spartan Books, New York, pp 1–14
- Medina Serrano JC, Papakyriakopoulos O, Hegelich S (2020) Exploring political ad libraries for online advertising transparency: lessons from Germany and the 2019 European elections. In: *International conference on social media and society*, pp 111–21.
- Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A (2019) A survey on bias and fairness in machine learning. *arXiv Preprint arXiv 1908.09635*
- Meira SR, Buregio VA, Nascimento LM, Figueiredo E, Neto M, Encarnação B, Garcia VC (2011) The emerging web of social machines. In: *2011 IEEE 35th annual computer software and applications conference*. IEEE, 26–27
- Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L (2016) The ethics of algorithms: mapping the debate. *Big Data Soc* 3(2):2053951716679679
- Mol A (2010) Actor-network theory: sensitive terms and enduring tensions. *Kölner Zeitschrift Für Soziologie Und Sozialpsychologie*. Sonderheft 50:253–269
- Mosier KL, Skitka LJ (1996) Human decision makers and automated decision aids: made for each other. In: *Automation and human performance: theory and applications*, pp 201–220
- Murray-Rust D, Davoust A, Papapanagiotou P, Manataki A, Van Kleek M, Shadbolt N, Robertson D. Towards executable representations of social machines. In: *International conference on theory and application of diagrams 2018 Jun 18*. Springer, Cham, pp 765–769
- Murray-Rust D, Robertson D (2015) Bootstrapping the next generation of social machines. In: *Crowdsourcing*. Springer, 53–71
- Murray-Rust D, Segolene T, Mark H, Owen G (2015) On wayfaring in social machines. In: *Proceedings of the 24th international conference on world wide web*, pp 1143–8. ACM
- Newell A, Simon HA et al (1972) *Human problem solving*. Prentice-Halls, Englewood Cliff
- Norman DA, Stappers PJ (2015) *DesignX: complex sociotechnical systems*. *She Ji J Des Econ Innov* 1(2):83–106
- Novikov DA (2015) *Cybernetics: from past to future*, vol 47. Springer, Berlin
- Palermos SO (2017) Social machines: a philosophical engineering. *Phenomenol Cogn Sci* 16(5):953–978
- Papakyriakopoulos O, Serrano JCM, Hegelich S (2020) The Spread of Covid-19 conspiracy theories on social media and the effect of content moderation. In: *The Harvard Kennedy School (HKS) misinformation review*, vol 1
- Papapanagiotou P, Davoust A, Murray-Rust D, Manataki A, Van Kleek M, Shadbolt N, Robertson D (2018) Social machines for all. In: *Proceedings of the 17th international conference on autonomous agents and multiagent systems. international foundation for autonomous agents; multiagent systems*, pp 1208–1212
- Pariser E (2011) *The filter bubble: how the new personalized web is changing what we read and how we think*. Penguin, New York
- Pasquale F (2015) *The black box society: the secret algorithms that control money and information*. Harvard University Press, Cambridge
- Piaget J (1947) *The psychology of intelligence*. Routledge, New York
- Rahwan I, Cebrian M, Obradovich N, Bongard J, Bonnefon J-F, Breazeal C, Crandall JW et al (2019) Machine behaviour. *Nature* 568(7753):477–486
- Rainie L, Smith A, Schlozman KL, Brady H, Verba S (2012) Social media and political engagement. *Pew Internet Am Life Proj* 19:2–13
- Ratkiewicz J, Conover M, Meiss M, Gonçalves B, Flammini A, Menczer F (2011) Detecting and tracking political abuse in social media. In: *Fifth international AAAI conference on weblogs and social media*
- Rentschler CA (2014) Rape culture and the feminist politics of social media. *Girlhood Stud* 7(1):65–82
- Ribeiro MH, Ottoni R, West R, Almeida VA, Meira Jr W (2020) Auditing radicalization pathways on youtube. In: *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pp 131–141
- Rosenblueth A, Wiener N, Bigelow J (1943) Behavior, purpose and teleology. *Philos Sci* 10(1):18–24
- Roush W (2005) Social machines: computing means connecting. *Technol Rev Manchester NH* 108(8):44
- Ruesch J, Bateson G, Pinsker EC, Combs G (2017) *Communication: the social matrix of psychiatry*. Routledge, New York
- Ruppert E, Isin E, Bigo D (2017) Data politics. *Big Data Soc* 4(2):2053951717717749
- Sandvig C, Hamilton K, Karahalios K, Langbort C (2014) Auditing algorithms: research methods for detecting discrimination on internet platforms. *Data Discrim Convert Crit Concerns Product Inquiry*. 22:4349–4357
- Schipper BC, Woo H (2018) Political awareness, microtargeting of voters, and negative electoral campaigning. <https://doi.org/10.2139/ssrn.2039122>
- Seaver N (2017) Algorithms as culture: some tactics for the ethnography of algorithmic systems. *Big Data Soc* 4(2):2053951717738104
- Seaver N (2019) Knowing algorithms. In: *digitalSTS: a field guide for science & technology studies*, pp 412
- Selbst AD, Danah B, Sorelle AF, Suresh V, Janet V (2019) Fairness and Abstraction in Sociotechnical Systems. In: *Proceedings of the conference on fairness, accountability, and transparency*, pp 59–68

- Shadbolt N, O'Hara K, De Roure D, Hall W (2019) The theory and practice of social machines. Springer, Berlin
- Shahrezayeh M, Orestis P, Juan CMS, Simon H (2019) Measuring the ease of communication in bipartite social endorsement networks: a proxy to study the dynamics of political polarization. In: Proceedings of the 10th international conference on social media and society. ACM, pp 158–165
- Shneiderman B, Plaisant C (2010) Designing the user interface: strategies for effective human-computer interaction. Pearson Education India, Chennai
- Smart PR, Shadbolt NR (2015) Social machines. In: Encyclopedia of Information science and technology, 3rd edn. IGI Global, pp 6855–6862
- Smart P, Simperl E, Shadbolt N (2014) A taxonomic framework for social machines. In: Miorandi D, Maltese V, Rovatsos M, Nijholt A, Stewart J (eds) Social collective intelligence. Computational Social Sciences. Springer, Cham, pp 51–85. https://doi.org/10.1007/978-3-319-08681-1_3
- Stieglitz S, Dang-Xuan L (2013) Social Media and Political Communication: A Social Media Analytics Framework. *Soc Netw Anal Min* 3(4):1277–1291
- Taylor L (2017) What is data justice? The case for connecting digital rights and freedoms globally. *Big Data Soc* 4(2):2053951717736335
- Tufekci Z, Wilson C (2012) Social media and the decision to participate in political protest: observations from Tahrir square. *J Commun* 62(2):363–379
- U-Directive E (2016) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Off J Eur Union* L119(May):1–88
- Vicario D, Michela AB, Zollo F, Petroni F, Scala A, Guido Caldarelli H, Stanley E, Quattrociocchi W (2016) The spreading of misinformation online. *Proc Natl Acad Sci* 113(3):554–559
- Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. *Science* 359(6380):1146–1151
- Vygotski L (2012) Thought and language. MIT Press, Cambridge
- Wachter S, Mittelstadt B, Russell C (2017) Counterfactual explanations without opening the black box: automated decisions and the Gdpr. *Harv JL Tech* 31:841
- Wiener N (2019) Cybernetics or control and communication in the animal and the machine. MIT Press, Cambridge
- Yang JA (2016) Effects of popularity-based news recommendations ('most-viewed') on users' exposure to online news. *Media Psychol* 19(2):243–271
- Zarsky T (2016) The trouble with algorithmic decisions: an analytic road map to examine efficiency and fairness in automated and opaque decision making. *Sci Technol Hum Values* 41(1):118–132
- Ziewitz M (2016) Governing algorithms: myth, mess, and methods. *Sci Technol Hum Values* 41(1):3–16
- Zuckerman E (2014) New media, new civics? *Policy Internet* 6(2):151–168
- Zuckerman E (2019) Beyond the Vast Wasteland: Briefing Congresspeople for the Aspen Institute. My hearts in Accra. <http://www.ethanzuckerman.com/blog/2019/07/31/beyond-the-vast-wasteland-briefing-congresspeople-for-the-aspen-institute/> Accessed 10 Mar 2020
- Zuiderveen Borgesius F, Möller J, Kruikemeier S, Ó Fathaigh R, Irion K, Dobber T, Bodo B, de Vreese CH (2018) Online political microtargeting: promises and threats for democracy. *Utrecht Law Rev* 14(1):82–96

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